



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)</b>	<b>A2</b>	(11) International Publication Number: <b>WO 00/20601</b>
		(43) International Publication Date: 13 April 2000 (13.04.00)

(21) International Application Number: PCT/US99/22886

(22) International Filing Date: 1 October 1999 (01.10.99)

## (30) Priority Data:

60/102,748	2 October 1998 (02.10.98)	US
60/123,810	11 March 1999 (11.03.99)	US
60/139,650	17 June 1999 (17.06.99)	US

(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).

## (72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).

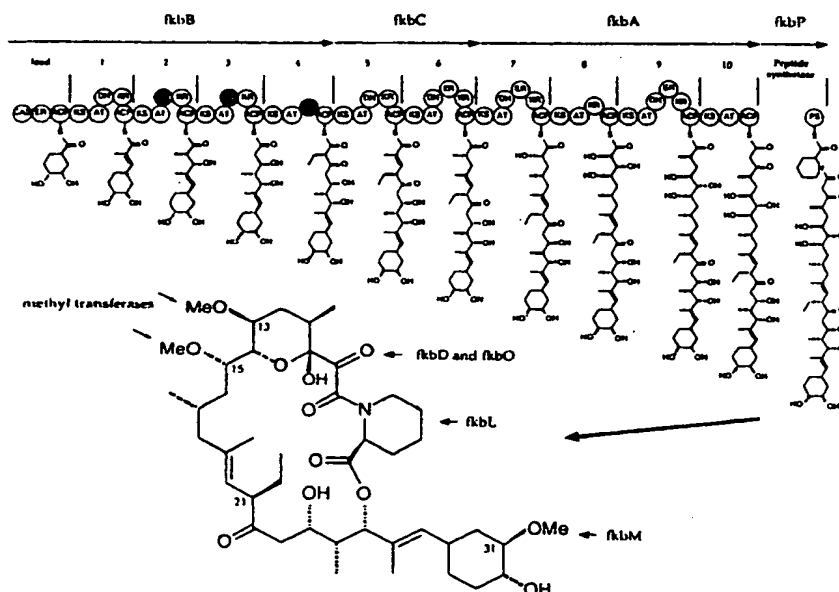
(74) Agents: FAVORITO, Carolyn et al.; Morrison &amp; Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).

(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SI, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

## Published

Without international search report and to be republished upon receipt of that report.

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



## (57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NI	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA  
CONSTRUCTS THEREFOR

5

Field of the Invention

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to  
10 compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fields of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

Background of the Invention

15 Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes. Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline,  
20 erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

25 This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either to produce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce  
30 molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5,063,155; 5,098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5,843,718; and Fu *et al.*, 1994, *Biochemistry* 33: 9321-9326; McDaniel *et al.*, 1993, *Science* 262: 1546-1550; and Rohr, 1995, *Angew. Chem. Int. Ed. Engl.* 34(8): 881-888,  
35 each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal peptide synthetase (NRPS). Two major types of PKS enzymes are known; these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as *eryAI*, *eryAII*, and *eryAIII*. See Caffrey *et al.*, 1992, *FEBS Letters* 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS<sup>Q</sup>, where the superscript letter is the abbreviation for the amino acid, glutamine, that is present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module



incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

5           The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior  
10       module, optionally performing one or more additional functions, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS, AT, ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender  
15       module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

20           Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-  
25       carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender  
30       module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom, producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS, AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, a DH domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activities, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes; these enzymes add carbohydrate groups or methyl groups, or make other modifications, i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of erythromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypeptides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those taken from other sources. A genetically engineered

PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence  
5 alignments also have revealed linker regions between the catalytic domains and at the N- and C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and between the individual polypeptides that comprise the PKS complex. One  
10 can thus view the linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT  
15 replacement, one can thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that  
20 known polyketides can be produced more effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes.  
25 The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present  
30 invention helps meet the need for such compounds as well.

#### Summary of the Invention

In one embodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention  
35 include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3,

pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

The invention also provides isolated nucleic acids that encode all or part of one or more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-506 or rapamycin.

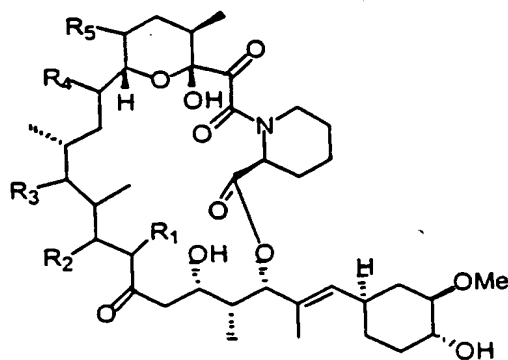
In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be

used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppression activities.

Thus, the invention provides polyketides having the structure:



wherein, R<sub>1</sub> is hydrogen, methyl, ethyl, or allyl; R<sub>2</sub> is hydrogen or hydroxyl, provided that when R<sub>2</sub> is hydrogen, there is a double bond between C-20 and C-19; R<sub>3</sub> is hydrogen

or hydroxyl; R<sub>4</sub> is methoxyl, hydrogen, methyl, or ethyl; and R<sub>5</sub> is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

5 In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

These and other embodiments and aspects of the invention will be more fully  
10 understood after consideration of the attached Drawings and their brief description below, together with the detailed description, examples, and claims that follow.

### Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line  
15 provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbc*. Immediately under the third line are numbered segments showing where the loading  
20 module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various cosmids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the  
25 peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3,  
30 and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting  
35 at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the

methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chain on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol.* 39:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fk bD*, *fk bM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fk bN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fk bQ* (a type II thioesterase, which can increase polyketide production levels), and *fk bS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

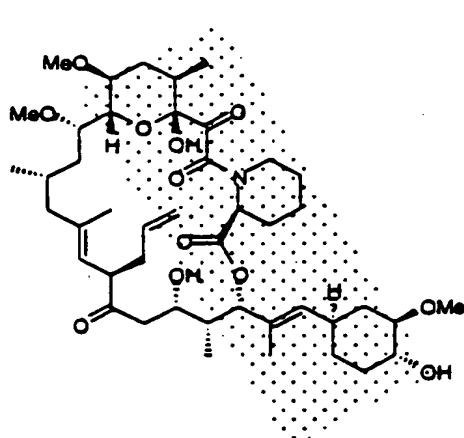
Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

#### Detailed Description of the Invention

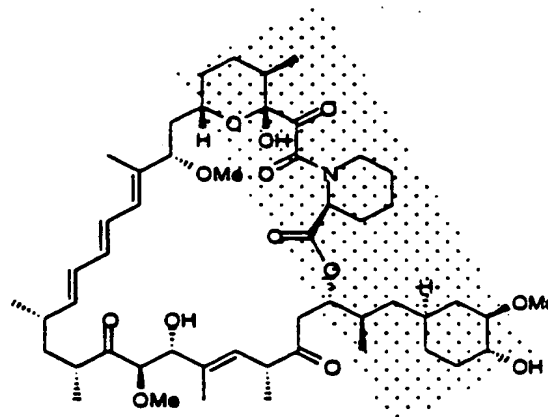
Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such

methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see Holt *et al.*, 1993, *JACS* 115:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart,  
 5 kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the unapproved use of tacrolimus for other conditions, including alopecia  
 10 universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.



FK-506



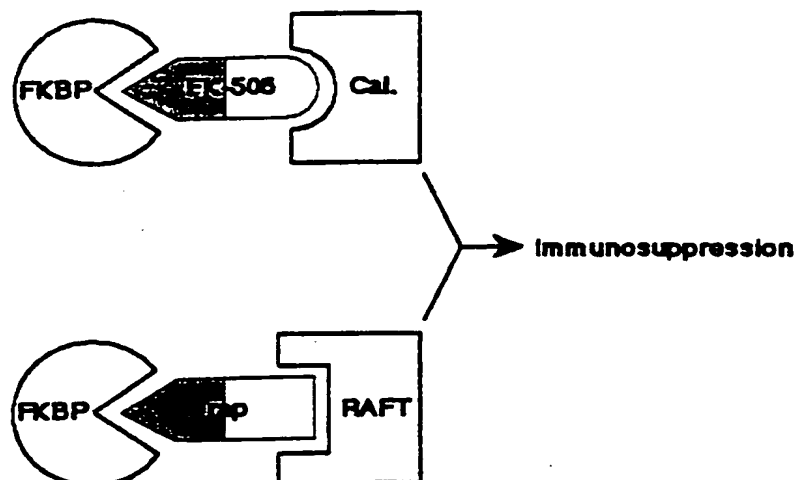
Rapamycin

FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

These compounds act through initial formation of an intermediate complex with  
 20 protein "immunophilins" known as FKBP (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules,  
 25 known as the "FKBP-binding domain" (as generally but not precisely indicated by the



stippled regions in the structures above). The FK-506-FKBP complex then binds calcineurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506, FK-520, and rapamycin. Modifications in the effector domains of FK-506, FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont *et al.*, 1992, *Journal of Experimental Medicine* 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

In addition to immunosuppressive activity, FK-520, FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15:

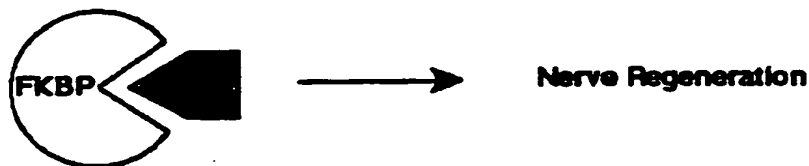
7509-7516; and Steiner *et al.*, 1997, *Proc. National Academy of Science* 94: 2019-2024.

Further, the restored central and peripheral neurons appear to be functional.

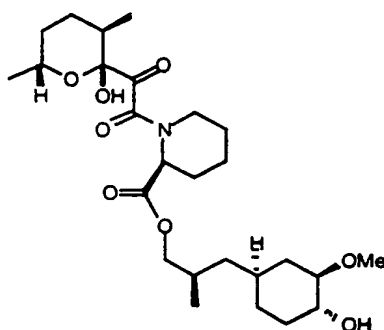
Compared to protein neurotrophic molecules (BNDF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects.

Finally, whereas protein neurotrophins often show effects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine* 3: 421-428.



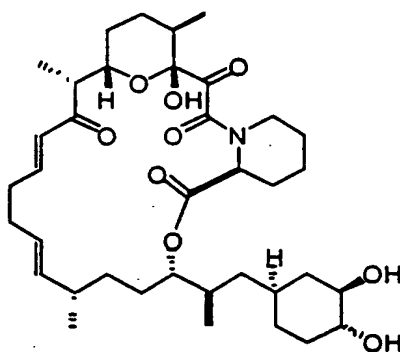
Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne *et al.*, 1993, *Journal of Molecular Biology* 229: 105-124.). Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.



"FKBP binding domain"

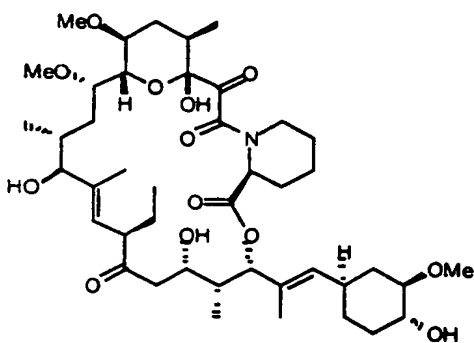
There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineurin or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

- 5 Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr *et al.*, 1996, *The Journal of Antibiotics* 49: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

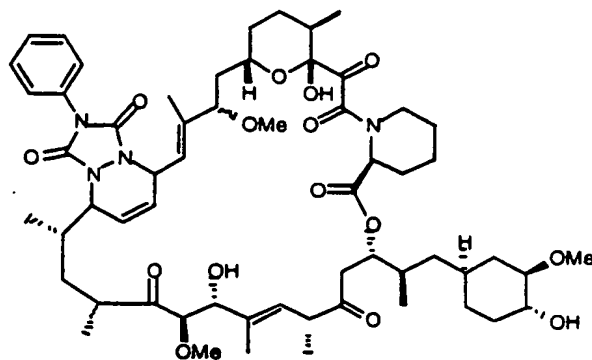


Antascomycin A

- 10 Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited,
- 15 some useful chemically modified analogs exist. The FK-520 analog L-685,818 ( $ED_{50}$  = 0.7 nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 ( $IC_{50}$  = 12.5 nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications* 192: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).

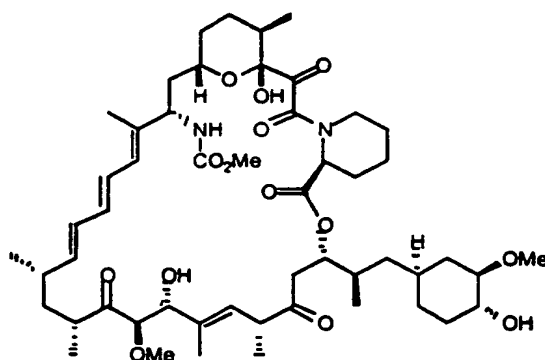


L-685,818



WAY-124,466

One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology* 2: 471-481). One of the best compounds, 1, below, shows complete loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.



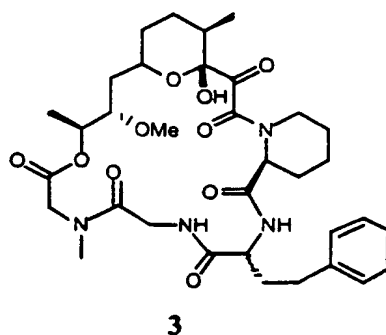
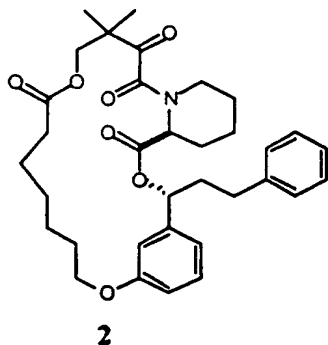
1

10

There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society* 115: 9925-9938); the best analog, 2, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the FKBP binding domain were tethered by a tripeptide to give analog 3, below, which binds

15

to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.



5 In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand  
10 restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

From the above description, two general approaches towards the design of non-  
15 immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by computational methods, and the analogs closely resemble parent molecules that have  
20 proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for production of the numerous compounds needed for such  
25 interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

The present invention provides useful methods and reagents related to the first  
30 approach, but with significant advantages. The invention provides recombinant PKS

genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures *via* genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin); similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab initio* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolism and pharmacokinetics of tacrolimus has been extensively studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%, (range 5 to 65%). The volume of distribution (V<sub>0</sub>D) based on plasma is 5 to 65 L per kg of body weight (L/kg), and is much higher than the V<sub>0</sub>D based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha<sub>1</sub>-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels. Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent *et al.*, 1992, *In vitro* metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, *Arch. Biochem. Biophys.* 294: 454-460; Iwasaki *et al.*, 1993, Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, *Drug Metabolism & Disposition* 21: 971-977; Shiraga *et al.*, 1994, Metabolism of FK-506, a

potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, *Biochem. Pharmacol.* 47: 727-735; and Iwasaki *et al.*, 1995, Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506, *Drug Metabolism & Disposition* 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as  
5 important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy)  
10 compounds undergo cyclizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII, was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation.  
15 Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-506, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).  
20

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important  
25 biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-506 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as  
30 tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the  
35 compounds.



These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and/or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood  
5 can vary widely depending on interactions with drugs that induce or inhibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawa □ US, Rev 4/97, Rec 6/97). Also important are the drugs that induce P450 3A  
10 (e.g. Dexamethasone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant  
15 adverse effects resulting from the use of FK-506 and are believed to be similar for FK-520. Because these effects appear to occur primarily by the same mechanism as the immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose related and correlates with high blood levels of the drug (Prograf package insert,  
20 Fujisawa □ US, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by the present invention should be more controllable, the incidence of toxicity should be significantly decreased with the 13-desmethoxy analogs. Some reports show that certain FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher  
25 therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant  
30 proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the *fkfA*, *fkfB*, *fkfC*, and *fkfP* gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the *fkfD* gene product and that is oxidized by the *fkfO* gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the *fkfM* gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded by the *fkfG* gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. *ascomyceticus* recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful FK-520-related compound merely as a result of inactivation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and

functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCos™ vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of genomic DNA was partially digested with 4 units of *Sau*3A I for 20 min. in a reaction volume of 1 mL, and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

Based on recently published sequence from the FK-506 cluster (Motamedi and Shafiee, 1998, *Eur. J. Biochem.* 256: 528), a probe for the *fkfO* gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two *Eco*RI fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial digestion with *Sau*3AI, gel purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new *fkfM*

probe isolated using DNA from ATCC 14891. A probe representing the *fkpP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional cosmids, pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA, USA. The complete nucleotide sequence of the coding sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated *fkpB*, *fkpC*, *fkpA*, and *fkpP*. The *fkpB* open reading frame encodes the loading module and the first four extender modules of the PKS. The *fkpC* open reading frame encodes extender modules five and six of the PKS. The *fkpA* open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The *fkpP* open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

	<u>Nucleotides</u>	<u>Gene or Domain</u>
	complement (412 - 1836)	<i>fkpW</i>
	complement (2020 - 3579)	<i>fkpV</i>
30	complement (3969 - 4496)	<i>fkpR2</i>
	complement (4595 - 5488)	<i>fkpR1</i>
	5601 - 6818	<i>fkpE</i>
	6808 - 8052	<i>fkpF</i>
	8156 - 8824	<i>fkpG</i>
35	complement (9122 - 9883)	<i>fkpH</i>
	complement (9894 - 10994)	<i>fkpI</i>
	complement (10987 - 11247)	<i>fkpJ</i>
	complement (11244 - 12092)	<i>fkpK</i>
	complement (12113 - 13150)	<i>fkpL</i>
40	complement (13212 - 23988)	<i>fkpC</i>

	complement (23992 - 46573)	<i>fkfB</i>
	46754 - 47788	<i>fkfO</i>
	47785 - 52272	<i>fkfP</i>
	52275 - 71465	<i>fkfA</i>
5	71462 - 72628	<i>fkfD</i>
	72625 - 73407	<i>fkfM</i>
	complement (73460 - 76202)	<i>fkfN</i>
	complement (76336 - 77080)	<i>fkfQ</i>
	complement (77076 - 77535)	<i>fkfS</i>
10	complement (44974 - 46573)	CoA ligase of loading domain
	complement (43777 - 44629)	ER of loading domain
	complement (43144 - 43660)	ACP of loading domain
	complement (41842 - 43093)	KS of extender module 1 (KS1)
	complement (40609 - 41842)	AT1
15	complement (39442 - 40609)	DH1
	complement (38677 - 39307)	KR1
	complement (38371 - 38581)	ACP1
	complement (37145 - 38296)	KS2
	complement (35749 - 37144)	AT2
20	complement (34606 - 35749)	DH2 (inactive)
	complement (33823 - 34480)	KR2
	complement (33505 - 33715)	ACP2
	complement (32185 - 33439)	KS3
	complement (31018 - 32185)	AT3
25	complement (29869 - 31018)	DH3 (inactive)
	complement (29092 - 29740)	KR3
	complement (28750 - 28960)	ACP3
	complement (27430 - 28684)	KS4
	complement (26146 - 27430)	AT4
30	complement (24997 - 26146)	DH4 (inactive)
	complement (24163 - 24373)	ACP4
	complement (22653 - 23892)	KS5
	complement (21420 - 22653)	AT5
	complement (20241 - 21420)	DH5
35	complement (19464 - 20097)	KR5
	complement (19116 - 19326)	ACP5
	complement (17820 - 19053)	KS6
	complement (16587 - 17820)	AT6
	complement (15438 - 16587)	DH6
40	complement (14517 - 15294)	ER6
	complement (13761 - 14394)	KR6
	complement (13452 - 13662)	ACP6
	52362 - 53576	KS7
	53577 - 54716	AT7
45	54717 - 55871	DH7
	56019 - 56819	ER7
	56943 - 57575	KR7
	57710 - 57920	ACP7
	57990 - 59243	KS8
50	59244 - 60398	AT8
	60399 - 61412	DH8 (inactive)
	61548 - 62180	KR8

62328 - 62537  
 62598 - 63854  
 63855 - 65084  
 65085 - 66254  
 5 66399 - 67175  
 67299 - 67931  
 68094 - 68303  
 68397 - 69653  
 69654 - 70985  
 10 71064 - 71273

ACP8  
 KS9  
 AT9  
 DH9  
 ER9  
 KR9  
 ACP9  
 KS10  
 AT10  
 ACP10

1 GATCTCAGGC ATGAAGTCCT CCAGGCGAGG CGCCGAGGTG GTGAACACCT CGCCGCTGCT  
 61 TGTACGGACC ACTTCAGTCA GCGGCGATTG CGGAACCAAG TCATCCGGAA TAAAGGGCGG  
 121 TTACAAGATC CTCACATTGC GCGACCGCCA GCATACGCTG AGTTGCCTCA GAGGCAAACC  
 15 181 GAAAGGGCGC GGGCGGTCCG CACCAGGGCG GAGTACGCGA CGAGAGTGGC GCACCCGCGC  
 241 ACCGTCACCT CTCTCCCCCG CCGGCGGGAT GCCCGGCGTG ACACGGTTGG GCTCTCCTCG  
 301 ACGCTGAACA CCCGCGCGGT GTGGCGTCGG GGACACCGCC TGGCATCGGC CGGGTGACGG  
 361 TACGGGGAGG GCGTACGGCG GCCGTGGCTC GTGCTCACGG CCGCCGGGCG GTCATCCGTC  
 421 GAGACGGCAC TCGGCGAGCA GGGACGCCTG GTCGGCACCT GCGGGCCGGA CGACCGTGTG  
 20 481 GTTCGCGGGC GGGCGGTGGC CGGTGGTGAG CCAGCTCTCC AGGGCGGTGA AGGCTGAGCG  
 541 GTGACACGGC AGCAAAGGCC GGAGTCGGTC GGGGAAGGTG TCGACGAGGG CGTCGGTGTG  
 601 CGTGCCGTCC TCGATGCGGT AGTAGCGGTA CCGGCCGCCA GGCCGCTGCC GGACATACGC  
 661 GCGTACACGT CGGAGCCCGG GCGGCAGGCA GCAGCACGTC GAGAGTGCCT GGATGGTGAT  
 721 CAGCGGCTTG CCGATACGAC CGGTCAACGC GATGCGTTCC ACGGCCGCGT GGACGCGGA  
 25 781 GGAGCGGGTG GCGTAGTCGT AGTCGGCATC GCAGCCCGGG ACCGTCCCCG GGGCGCAATA  
 841 CGGTGIGCCG GCTTCCTTCT CCCCATCGAA GCCGGGGTCG AACTCCTCGC GGTAGACGCG  
 901 CTGCGTCAGA TCCCAGTAGA CCTCGTGGTG GTACGGCCAC AAGAACTCGG AGTCGGCCCG  
 961 GAACCCGGCG CCGAGCAGCG CCTCGCGCGC CTGGCCGGCT GCGGGGCCCG CTGCCGCGTA  
 1021 GGTGGGGTAG TCGCGCAGGG CCGCCGGCAG GAAGGTGAAG AGGTTGGGAC CCTCCGCGCG  
 30 1081 CCACAGGGTG CCTTCCAGT CCACTCCTCC GTCGTACAGC TCGGGATGGT TCTCCAGCTG  
 1141 CCAGCGCACG AGGTAGCCGC CGTTGGACAT CCCGGTGACC AGGGTGCGCT CGAGCGGCCG  
 1201 GTGGTAGCGC TGGGCGACCG ACGCGCGGGC GGCCCGGGTC AGCTGGGTGA GGCGGGTGTT  
 1261 CCACTCGGCG ACGGCGTCGC CCGGCCGGGA GCCATCACGG TAGAACGCGG GGCCTGTT  
 1321 GCCCTTGTCG GTGGCGGCGT AGGCGTAACC GCGGGCGAGC ACCCAGTCGG CGATGGCCCC  
 35 1381 GTCGTTGGCG TACTGCTCGC GGTACCAGG GGTGCCGGCC ACGACAGGC CACCGTTCCA  
 1441 GCGGTCGGGC AGCCGGATGA CGAACTGGGC GTCGTGGTTC CACCCGTGGT TGGTGTGGT  
 1501 GGTGGAGGTG TCGGGGAAGT AGCCGTGAT CTGGATCCCG GGCCTCCGG TGGAGTGGC  
 1561 CAGGTTCTTG GCGGTCAGCC CTGCCCAGTC CGCCGGGTCG GTGTGGCCGG TGGCCGCCGT  
 1621 TCCCGCCGTG GTCAGCTCGT CCAGGCAGTC GGCCTGCTGA CGTGCCGCCG CCGGGACACG  
 40 1681 CAGCTGGGAC AGACGGGCGC AGTGACCGTC CCGGGCATCG GGAGCAGGCC GGGCGGTGGC  
 1741 CGGTGAGGGG AGCAGGACGG CCACTGCGGC CAGGGTGAGA GCGCCGAGGC GCGTGTCTCT  
 1801 TCTCGGGGCC CGTCCGACAC CGAGGGGCAG AACCATGGAG AGCCTCCAGA CGTGCGGATG  
 1861 GATGACGGAC TGGAGGCTAG GTGCGCACG GTGGAGACGA ACATGGGTGC GCCCGCATG  
 1921 ACTGAGGCC CACAGAGGTG GGCCGCGCC ATGACGGGCG CCGGACCGCG GCGCTCCGG  
 45 1981 GGCGGIGCCC GCGGCCGCCA CCGGTTCCGG GTCCCGGGT CAGGGACAGG TGTCGTTCCG  
 2041 GACCGTGAAG TAGCCGGTCG GCGACTCTTT CAAGGTGGTC GTGACGAAGG TGTTGTACAG  
 2101 GCCCATGTTT TGGCCGGAGC CCTTGGCGTA GGTGTAACCG GCGTCTGTCG TGGCGCGGCC  
 2161 CGCCTGGACG TGAGCGTAGT TGCCGGCGGT CCAGCAGACG GCCGTGGCAC GCGTGTCTG  
 2221 CGCGGTGACC GCGCCCGAGA GCGGTCCGGC CTTGCCGTCC GCGTCCCGGG CCGCGACCGC  
 50 2281 GTAGGTGTGC GATGTGCCCG CCCTCAGGCC GGTGTCCGTG TACGACGTCG TGGCGGACGT  
 2341 GGTGATCTGG GCACCGTCGC GGTGGACGGC GTAGTCGGTG GCGCCGTCGA CCGGTTTCCA  
 2401 GGTGAGGCTG ATGGTGGTGT CCGTGGCGCC GGTGGCGGCC AGGCCGGACG GAGCGGGCAG  
 2461 CGAACCGGG TCGGAGGCGG ATCCGCTCAG GCCGAAGAAC TGCGTGATCC AGTAGTGGGA  
 2521 CACAGTCGAG TCCAGGAAGT AGCGGCGGCC GGTGCTGCC CACTGTGTCG CTCGGTGCC  
 55 2581 GGGATCGACC GGGGTGCCGT GCCGATGCC CCGCACCCGG TTACCTCCA CCGCCACCGA  
 2641 TCCGTCCGCG GCCAGGTA CTTCTGTGCC GGTGGAGTTC GGGCCGATCA CCGAGGTACG  
 2701 GTCCGGCGTC TGGGACACGC CGTGACAGC GGTCCACTGG TCGCGCAACT CGTCGGCGTT  
 2761 GCGCGGCGCG ACGGTGGTGT CTTGTGCGCC GTGCCAGATG GCCACGCGCG GCCACGGGCC  
 2821 CGACCACGAG GGGTAGCCGT CACGGACCCG CCGCGCCAC TGGTCCGCGG TCAGGTCCGT  
 60 2881 CCGGGGGTTC ATGCACAGGT ACGCGCTGCT GACGTCCGTG GCACAGCCGA AGGGCAGGCC  
 2941 GGCGACGACC GCGCCGGCCT GGAAGACGTC CCGATAGGTG GCGAGCATCA CCGACGTCAT

3001 GGCACCGCCG GCGGACAGCC CGGTGATGTA GGTGCGCTGG GGGTCCGCGC CGTAGGCGGA  
 3061 GACGGTGTGA GCGGCCATCT GCCGGATCGA CGCGGCTTCG CCCTGGCCCC TGCGGTTGTC  
 3121 GCTGCTCTGG AACCAGTTGA AGCACCTGTT CGCGTTGTTC GACGACGTGG TCTCGGCGAA  
 3181 CACGAGCAGG AAGCCATAGC GGTCCGCGAA TGAGAGCAGG CCGGAGTTGT CGGCGTAGCC  
 5 3241 CTGGGCGTCC TGGGTGCAAC CGTGCAGGGC GAACACCACC GCCGGCTCCG CGGGCAGGGA  
 3301 CGCGGGCCGG TAGACGTACA TGTTTCAGCCG GCCCGGGTTC GTGCCGAAGT CCGCGACCTC  
 3361 GGTCAGGTCC GCCTTGGTCA GACCGGGCTT GGCCAGGCCC GCCGCGGCGT GGGCCGTCCG  
 3421 CGCCGGGCGG AGCAGGGCCG CTCCGAGTAC GAGGGCCACG ACGGCCACGA GACGGGAGAG  
 3481 CACCCCCCGC CGTCCCGGAC GCGACAACGA CCCGACCGGC GCGGAGGAGG AGAGGGGGA  
 10 3541 CACGCGGGTG AGGATTCCCC GGAACGGCGG CGGCTGCATG GCGGCTCCCT CATGTCTGTG  
 3601 GGGGGGACAC GGAGGGCTCC CTGACGTCGA TCAGTGGGAG CGCCCCGGTG CCCGGCACCG  
 3661 TAGGGGTGGT TCAACCCGCA ACGGTATGGC CCGGAGCACC ACACCCCGCA CCGCGCGATG  
 3721 TGCGCCCCGA CGGATTGTGT CGCCTTGCGG AATCTGATAC CCGGACGCGA CGAACGCCCC  
 3781 ACCCGACACG GGTAGGGCGT CATGGTGTCC GACTCGGCCG GTCGGCCTTG CCTGCCCTGG  
 15 3841 ACGGACCGGG CGTCGGCGGA CCGGGCGTCG GCGGGCTGGG CCGTATGGCG GCCGAGGACG  
 3901 CCAGCCGCGT GGGGCGGCGG CGCCCAAGTG CAGTACGCCG ACCGTGGCCG CCGGAGGGCG  
 3961 CGGACCGGTC AGTGCAGTCC CGCGGCCCTG CGGGACCGCT CGTCCCAGAC GGGTTCCACC  
 4021 GCGGCGAACC GGGGTCCGTG TCCGCGGCGG TAGACCATCA GTGTCCGCTC GAAGGTGATG  
 4081 ACGATGACAC CGTCCTGGTT GTAGCCGATG GTGCGCACGC TGATGATGCC TACGTCAGGT  
 20 4141 CGGCTGGCGG ACTCCCCGGT GTTCAGGACC TCGGACTGCG AGTAGATGGT GTCGCCCTCG  
 4201 AAGACCGGGT TCGGCAGCCT GACCCGGTCC CAGCCGAGGT TGGCCATCAC ATGCTGGGAG  
 4261 ATGTCCGGTGA CGCTCTGCCC GGTGACCAGG GCGAGGGTGA AGGTGGAGTC CACCGAGGGC  
 4321 TTGCCCCAGG TGGTSCCCGC CGAGTAGTGG CGGTCTGAAGT GCAGCGGCGC GGTGTTCTGC  
 4381 GTCAGGAGCG TGAGCCAGGA GTTGTGCGTC TCCAGGACCG TGCGGCCAG GGGGTGGCGG  
 25 4441 TACACGTCGC CGGTGGTGAA GTCTCGAAG TAGCGGCCCT GCCAGCCCTC GACCACAGCG  
 4501 GTGCGGGTGG CGTCCTGGTC CCGGTTCTCA GTCGTCATGG CGCTCATTCT GGGAAAGTCCC  
 4561 CCGTCCGCTG TGAATGCCG AACCTTCACC GGGCTCATAC GTGCGGCGCA TGAGCCCTGG  
 4621 ACCGTACGTA GTCGTAGAAC CTCGCCACCA CTGGCGCGCG TGGTCTCTCC GCGAGTGTGA  
 4681 CCACGTCGAC CGTGCGCCCG GCCTGCGGGT CGTCGAGCGG CACGGCGACG CGCTGGTCAC  
 30 4741 CCGGCCCCGA CCGGCTGCGG GTGAGGGGGG CGACGGCCAC ACCGAGCCG CCGGCGACCA  
 4801 GGGCCCGCAG CGTGCTCAGC TCGGTGCTCT CCAGGACGAC CCGCGGCACG AATCCGGCCG  
 4861 CCGCGCACAG CCGGTCCGTG ATCTGGCGCA GTCCGAAGAC CCGTCCAGT GCCACGAACG  
 4921 CCTCATCGGC CAGCTCCGCG GTCCGCAACC GCGGCGCTCT GGCCAGCCGG TGTCCGGGTG  
 4981 GGACGAGCAG GCACAGTGCC TCGTCCCGCA GTGGTGTCCA CTCCACATCG TCCCCGGCGG  
 35 5041 GTCGTGGGCT GGTGAGCCCC AGGTCCAGCC TGCTGTTGCG GACGTCTGCG ACCACGGCGT  
 5101 CCGGCGCGTC CCGCGCAGT TCGAAGGTGG TGCCGGGAGC CAGCGCGCGG TACCCGCGGA  
 5161 GGAGGTGCGG CACCAGCCAG GTGCCGTAGG AGTGCAGGAA ACCCAGTGCC ACGGTGCCCG  
 5221 TGTCGGGGTC GATCAGGGCG GTGATGCGCT GCTCGGCGCC GGAGACCTCA CTGATCGCGC  
 5281 GCAGGGCGTG GCGCGGAAG ACCTCGCCGT ACTTGTGAG CCGGAGCCGG TTCTGGTGCC  
 40 5341 GGTCGAACAG CCGCACGCCC ACTCGTCGCT CCAGCCGCGG GATGGCCCTG GACAGGGTCG  
 5401 GCTGGGAGAT GTTGAGCCGT TCCGCGGTGA TCGTCACGTG CTCGTGCTCG GCCAAGGCCG  
 5461 TGAACCACTG CAACTCCCCT ATCTCCATGC AGGGACTATA CGTACCGGGC ATGCTCTGG  
 5521 CGAGGTTTCG TCATTTACA GCGGCGGGG GCGGCGCCAC AGTGAGTCTT CACCAACCAG  
 5581 GACCCCATGG GAGGGACCCC ATGTCCGAGC CGCATCCTCG CCCTGAACAG GAACGCCCCG  
 45 5641 CCGGGCCCTT GTCCGGTCTG CTCGTGGTTT CTTTGGAGCA GGCCGTCGCC GCTCCGTTCTG  
 5701 CCACCCGCCA CCTGGCGGAC CTGGGCGCCC GTGTCATCAA GATCGAACGC CCCGGCAGCG  
 5761 GCGACCTCGC CCGCGGCTAC GACCGCACGG TCGTGGCAT GTCCAGCCAC TTCGTCTGGC  
 5821 TGAACCGGGG GAAGGAGAGC GTCCAGTCTG ATGTGCGCTC GCCGGAGGGC AACCGGACCC  
 5881 TGCACGCCCT GGTGGACCGG CCGGATGTCC TGGTGCAGAA TCTGGCACCC GCGCCGCGG  
 50 5941 GCGGCGCTGG ATCGGCCACC AGGTCTCTCG GCGGAGCCAC CGAGGCTGAT CACCTGCGGA  
 6001 CATATCCGGC TACGGCAGTA CCGGCTGCTA CCGCGGACCG CAAGGCGTAC GACCTCCTGG  
 6061 TCCAGTGCGA AGCGGGGCTG GTCTCCATCA CCGGCACCCC CGAGACCCCG TCCAAGGTGG  
 6121 GCCTGTCCAT CGCGGACATC TGTGCGGGGA TGTACGCGTA CTCCGGCATC CTCACGGCCC  
 6181 TGCTGAAGCG GGCCCGCACC GGCCGGGGCT CGCAGTTGGA GGTCTCGATG CTCGAAGCCC  
 55 6241 TCGGTGAATG GATGGGATAC CCGGAGTACT ACACGCGCTA CCGGCGCACC GCTCCGGCCC  
 6301 GCGGCGGCGC CAGCCACGCG ACGATCGCCC CTACCGGCCC GTTCACCACG CCGGACGGG  
 6361 AGACGATCAA TCTCGGGCTC CAGAACGAGC GGGAGTGGG TCTCTCTGTC GGTGTCTGTC  
 6421 TACAACGCCC CCGTCTCTGC GACGACCCGC GCTTTTCCGG CAACGCCGAC CCGGTGGCGC  
 6481 ACCGACCGA GCTCGACGCC CTGGTGAGCG AGGTGACGG CACGCTCACC GCGGAGGAAC  
 60 6541 TGGTGGCGCG GCTGGAGGAG GCGTGCATCG CCTACGCACG CCAGCGCACC GTGCGGGAGT  
 6601 TCAGCGAACA CCCCCAACTG CGTGACCGTG GACGCTGGG TCCGTTTCAG AGCCCGGTCTG  
 6661 GTGCGCTGGA GGGCCTGATC CCCCCTGTC CTTTCCACGG CGAGCACCAG CCGGCGCTGG  
 6721 GCCGGGTCCC GGAGCTGGGC GAGCATACCG AGTCCGTCTT GCGGTGGCTG GCGGCGCCCC  
 6781 ACAGCGCCGA CCGCGAAGAG GCCGCGCATG CCGAATGAAC TCACCGGAGT CCTGATCTG

	6841	GCCGCCGTGT	TCCTGCTCGC	CGGCGTACGG	GGGCTGAACA	TGGGCCTGCT	CGCGCTGGTC
	6901	GCCACCTTTC	TGCTCGGGGT	GGTCGCACTC	GACCGAACGC	CGGACGAGGT	GCTGGCGGGT
	6961	TTCCCCGCGA	GCATGTTCCCT	GGTGCTGGTC	GCCGTCACGT	TCCTCTTCGG	GATCGCCCCG
5	7021	GTCAACGGCA	CGGTGGACTG	GCTGGTACGT	GTGCGGGTGC	GGGCGGTGGG	GGCCCGGGTG
	7081	GGAGCCGTCC	CCTGGGTGCT	CTTCGGCCTG	GCGGCACTGC	TCTGCGCGAC	AGGCGCGGCC
	7141	TCGCCCGCGG	CGGTGGCGAT	CGTGGCGCCG	ATCAGCGTCG	CGTTCGCCGT	CAGGCACCGC
	7201	ATCGATCCGC	TGTACGCCGG	ACTGATGGCG	GTGAACGGGG	CCGCAGCCGG	CAGTTTCGCC
	7261	CCCTCCGGGA	TCCTGGGCGG	CATCGTCCAC	TCGGCGCTGG	AGAAGAACCA	TCTGCCCGTC
10	7321	AGCGGCGGGC	TGCTCTTCGC	AGGCACCTTC	GCCTTCAACC	TGGCGGTTCG	CGCGGTJTCA
	7381	TGGCTCGTCC	TCGGGCGCAG	GCGCCTCGAA	CCACATGACC	TGGACGAGGA	CACCGATCCC
	7441	ACGGAAGGGG	ACCCGGCTTC	CCGCCCCGGC	GCGGAACACG	TGATGACGCT	GACCGCGATG
	7501	GCCGCGCTGG	TGCTGGGAAC	CACGGTCCTC	TCCCTGGACA	CCGGCTTCCT	GGCCCTCACC
	7561	TTGCGCGCGT	TGCTGGCGCT	GCTCTTCCCG	CGCACCTCCC	AGCAGGCCAC	CAAGGAGATC
15	7621	GCCTGGCCCC	TGGTGTGCTG	GGTATGCGGG	ATCGTGACCT	ACGTCCGCCCT	ACGTCCGCCCT
	7681	CTGGGCATCG	TGGACTCCCT	GGGGAAGATG	ATCGCGGCGA	TCGGCACCCC	GCTGCTGGCC
	7741	GCCCTGGTGA	TCTGCTACGT	GGGCGGTGTC	GTCTCGGCCT	TCGCTCGAC	CACCGGGATC
	7801	CTCGGTGCCC	TGATGCCGCT	GTCCGAGCCG	TTCTGAAGT	CCGGTGCCAT	CGGGACGACC
	7861	GGCATGGTGA	TGGCCCTGGC	GGCCGCGGGC	ACCGTGGTGG	ACGCGAGTCC	CTTCTCCACC
20	7921	AATGGTGCTC	TGGTGGTGGC	CAACGCTCCC	GAGCGGCTGC	GGCCCGGCGT	GTACCAAGGG
	7981	TTGCTGTGGT	GGGGCGCCGG	GGTGTGCGCA	TGTGCTCCCG	CGGCGGCGCT	GGCGGCTTTC
	8041	GTGGTGGCGT	GAGCGCAGCG	GAGCGGGAAT	CCCCTGGAGC	CCGTTTCCCG	TGCTGTGTCG
	8101	CTGACGTAGC	GTCAAGTCCA	CGTGCCGGGC	GGGCAGTACG	CCTAGCATGT	CGGGCATGGC
	8161	TAATCAGATA	ACCCTGTCCG	ACACGCTGCT	CGCTTACGTA	CGGAAGGTGT	CCCTGCGCGA
25	8221	TGACGAGGTG	CTGAGCCGGC	TGCGCGCGCA	GACGGCCGAG	CTGCCGGGCG	GTGGCGTACT
	8281	GCCGGTGCAG	GCCGAGGAGG	GACAGTTCCT	CGAGTTCCTG	GTGCGGTTGA	CCGGCGCGCG
	8341	TCAGGTGCTG	GAGATCGGGA	CGTACACCGG	CTACAGCAGC	CTCTGCCTGG	CCCGCGGGAT
	8401	GGCGCCCGGG	GGCCGTGTGG	TGACGTGCGA	TGTCATGCCG	AAGTGCCCGC	AGGTGGGCGA
	8461	GCGGTACTGG	GAGGAGGCCG	GGGTTGCCGA	CCGGATCGAC	GTCCGGATCG	GCGACGJCCG
30	8521	GACCGTCCCT	ACCGGGCTGC	TCGACGAGGC	GGGCGCGGGG	CCGGAGTCGT	TCGACATGGT
	8581	GTTCAFCGAC	GCCGACAAGG	CCGGCTACCC	CGCTACTAC	GAGGCGGCGC	TGCCGCTGGT
	8641	ACGCCGCGGC	GGGCTGATCG	TCGTGCAAAA	CACGCTGTTC	TTCCGGCCGG	TGGCCGACGA
	8701	AGCGGTGCAG	GACCCGGACA	CGGTGCGGGT	ACGCGAACTC	AACGCGGCAC	TGCGCGACGA
	8761	CGACCGGGTG	GACCTGGCGA	TGCTGACGAC	GGCCGACGGC	GTCACCCCTG	TCCGGAAACG
35	8821	GTGACCGGGG	CGATGTCCGC	GGCGGTGAGC	GTCAGCGTCG	TCCGCGCGGG	TCCGCGGAG
	8881	GGCTCCAGAT	GCAGGCGTTC	GACGCCGGCG	GCGGAAGCGC	CCGCCACCTC	GGACACGCAG
	8941	GGGCGAGTCG	AGTCCGCGAA	GCCCGCGAAC	CGGTAGGCGA	TCTCCATCAT	GCGGTTGCGG
	9001	TCCGTACGCC	GGAAGTCCGC	CACCAGGTGC	GCCCCCGCGC	GGGCGCCCTG	GTCCGTGAGC
	9061	CAGTTCAGGA	TCGTGCGACC	GGCACCAGAA	GACACGACCC	GGCAGGACGT	GGCGAGCAGT
40	9121	TTCAGGTGCC	ACGTGCGAGC	CTTCTTCTCC	AGCAGGATGA	TGCCGACGGG	GCCGTGCGGG
	9181	CCGAAGCGGT	CGCCCATGGT	CGCCGACGAG	ACCTCATGGG	CGGGATCGGT	CGGACGCGC
	9241	GCAGGTGCGC	GTCGGAGTAG	TGCACGCCGG	TCGCGTTTAT	CTGGCTGGTC	CGCAGCGTCA
	9301	GTTCTCTGAC	GCGGCTGAGT	TCTCTCTCCC	CCGCGGGTGC	GATCGTTCAT	GAGAGGTGCA
	9361	GCGAGCGCAG	GAAGTCCTCG	TCGGGACCGG	AGTACGCCTC	CCGGGCTTGG	TCGCGCGCGA
45	9421	AACCCGCCTG	GTACATCAGG	CGGCGCCGAC	GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
	9481	ACTCCGGCAG	CGACAGGAGC	GTGGCCGCCT	GCTCGGCCGG	GTAGCACCAG	ACCTCGGGCA
	9541	GGTGGGAACG	CACCTCGGGA	CGCTCGGCGG	GCTGGTTCGT	GATGAACGCG	ATCGTGGTTC
	9601	GTGCGAAGTT	CAGCTCCGTG	GCGATCTCGC	GGACGGACTG	CGACTTCGGC	CCCATCCGA
	9661	TGCGGGCCAG	CACGAAGTAC	TCCGCCACAC	CGAGGCGTTC	CAGACGCTCC	CACGCGAGGT
50	9721	CGTGGTCGTT	CTTGCTCGCC	ACCGCCTGGA	GGATGCCGCG	GTGCTCGAGC	GTGGTGATCA
	9781	CCTCGCGGAT	CTCGTCCGGT	AGGACCACCT	CGTCTGCTTC	CAGCACGGTG	CCCCGCCACA
	9841	AGGTGTTGTC	CAGGTCCCAG	ACCAGACACT	TGACAATGGT	CATGGCTGTC	CTCTCAAGCC
	9901	GGGAGCGCCA	GCGCGTGCTG	GGCCAGCATC	ACCCGGCACA	TCTCGTGCTC	GCCCTCGATG
	9961	ATCTCCATGA	GCTTGGCGTC	CGGTTACGCC	CGTTGACGCA	CGTGTCCCTC	TCTCGCGCCT
55	10021	GCCGACGCGA	GCACCTGTGC	GGCGGTGCGG	GCCCCGGCGG	CGGCTCGTTC	GCGGCGGACG
	10081	TGCTTGGCCA	GGATCGTCGC	GGGCACCATC	TCGGGCGAGC	CCTCGTCCCA	GTGGTTCGCTG
	10141	GCGTACTCGC	ACACGCGGGC	CGCGATCTGC	TCCGCGGTCC	ACAGGTTCGG	GATGTGCCCG
	10201	GCGACGAGTT	GGTGGTTCGCC	GAGCGGCCGG	CCGAAGTCTG	CCCGGGTCCG	GGCGTGGGCC
	10261	ACCGCGGCGG	TGCGGCAGGC	CCGACGAGATC	CCGACGCGAG	CCGAGGCGAG	CGACTTGCGC
60	10321	CCGTAGGCGA	GTGACGCGGC	GACCGGATC	GGCAGTGACG	CGCCGGAGCC	GGCGAGGACC
	10381	GCGCCGCGCG	GCACACGCAC	TAGGTCCAGG	TGCAGATCGG	CGTGCCGCGG	GCGCGGCGAG
	10441	CCGACGCGCT	TCGGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTCGGCGGG	CACGACGACC
	10501	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACCA	GGTGGTCCGC	GTAGGCGGCG
	10561	GCAGTCGTCC	AGACCTTGTG	GCGCTCGACG	ACAGCGGTGT	CCCCGTGAG	CCGAACCCGC
	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGCCGAC	GGCCGCGAGT



	10681	TTCCTCGCTGG	TCAGCTCCTT	CAGGAAGGTC	GCCCGCTGAC	CGGCGTCGCC	GAGCCGCTGC
	10741	ACGGTCCACG	CGGCCATGCC	CTGCGACGTC	ATGACACTGC	GCAGCGAACT	GCAGAGGCTG
	10801	CCGACGTGTG	CGGTGAACTC	GCCGTTCTCC	CGGCTGCCGA	GTCCACAGACC	GCCGTGCTCG
	10861	GCCGCCACTT	CCGCGCAGAG	CAGGCCGTCG	GCGCCGAGCC	GGACGAGCAG	GTCGCGCGGC
5	10921	AGTTCGCCCG	ACGTGTCCCA	CTCGGCGGCC	CGGTACCCGA	CAAGGTCGGT	CAGCAGCGCG
	10981	TCACGCTCAG	GCATCGACGG	CCCGCAGCCG	GTGGACGAGT	GCGACCATGG	ACTCGACGGT
	11041	ACGGAAGTTC	GCGAGCTGGA	GGTCCGGGCC	GGCGATCGTG	ACGTGCAACG	TCTTCTCCAG
	11101	GTACACGACC	AGTTCCATCC	CGAACAGCGA	CGTGAGGCCG	CCCTCCGCGA	ACAGGTCGCG
	11161	GTCCACGGGC	CAGTCCGACC	TGGTCTTCGT	CTTGAGGAAC	GCGACCAACG	CGTCCGCGAC
10	11221	GGGGTCGTCC	TTGACGGGTG	CGGTCTATGAG	AACACCTTCT	CGTATTCTGTA	GAAGCCCGCG
	11281	CCGGTCTTCC	GGCCGTGGTG	TCCCTCGCGG	ACCTTGCCCA	GCAGCAGGTC	ACAGGGGCGG
	11341	CTGCGCTCGT	CGCCGGTGCG	TTTGTGCAGC	ACCCACAGCG	CGTCGACGAG	GTTGTCTGATG
	11401	CCGATCAGGT	CCGCGGTGCG	CAGCGGCCCG	GTCCGATGGC	CGAGGCACCC	CGTCATGAGC
	11461	GCGTCGACGT	CCTCGACGGA	CGCGGTGCC	TCCTGCACGA	TCCGCGCCGC	GTCGTTGATC
15	11521	ATCGGGTGGA	GCAGCGGCT	CGTGACGAAG	CCGGGCGCGT	CCCGGACGAC	GATCGGCTTG
	11581	CGCCGACGCG	CCGCGACGAG	CTCCCCGCGC	GCGGCCATGG	CCCTTCTACC	GATCCGGGGT
	11641	CCGCGGATCA	CCTCGACCGT	CGGGATCAGG	TACGACGGGT	TCATGAAGTG	CGTGCCGAGC
	11701	AGGTCTCTCG	GCCGGGCCAC	GGAGTCGGCC	AGTTCGTCAA	CCGGGATCGA	CGACGTGTTC
	11761	GTGATGACCG	GGATACCGGG	CGCCGCTGCC	GAGACCGTGG	CGAGTACCTC	CGCTTGACC
20	11821	TCCGGCTCCT	CGACGACGGC	CTCGATCACC	GCGGTGGCCG	TACCGATCGC	GGGACGCGCG
	11881	GACGTGGCCG	TCCGACGAC	ACCGGGTTCG	GCCCTCGGCG	GCCCGGCCAC	GAGTTGTGCC
	11941	GTCCGAGTT	CGGTGGCGAT	CCGCGCCGCG	CCGCGCGTAA	GGATCTCTC	GGACGTGTCG
	12001	ACGAGTGTCA	CCGGGACGCC	GTGGCGCAGC	GCGAGCGTGG	TGATGCGCGT	GCCCATCTACT
	12061	CCCGCGCCGA	GCACGATCAG	CTGGTGGTCC	ACGCTGTTTC	CTCCCTCCGG	GGTCACCATG
25	12121	GCAGCGAGTA	CGGGTCGAGG	ACGTCTTCCG	GGGTCGACCC	GATCGCGTCC	TTGCGGCCGA
	12181	GGCCGAGTTC	GTGCGCGAAG	CCGAGCAGCA	CGTCGAACGC	GATGTGGTCC	GCGAACGCGC
	12241	TGCCCCCTCGA	GTGAGGACG	CTCAGGCTGT	CCCGGTGGTC	CGCCGCGGTG	TCCGGTGCCG
	12301	CGCACAGGGC	CGCCAGCGAC	GGGCGAGCT	CGCGGTCCGG	CAGTTGCTGG	TACTCGCCCT
	12361	CGCGCGGGGC	CTGCCCGGGA	TGGTCGACGC	AGATGAACGC	GTCGTCGAGC	AGGGTCTTCG
30	12421	GCAGTTCGGT	CTTGCCCGGC	TCGTGCGGCG	CGATGGCGTT	CACATGCAGG	TGCGGCAGCC
	12481	GCGGCTCGGC	GGGACGACCC	GGCCCTTTGC	CCGAGGGCAC	CGAGGTGACG	GTGGACAGGA
	12541	CATCCGCGGC	GGCGGCGGCC	TCCGCCGGAT	CGGTACCTT	GACCGGCAGT	CCGAGGAACG
	12601	CGATGCGGTC	CGCGAACGAC	GCCGCGTGGC	CGGGGTCCGT	GTCGCTGACC	AGGATCCGCT
	12661	CGATGGGCAG	GACCCTGCTG	AGCGCGTGCG	CCTGGGTAC	CGCCTGTGCG	CCCGCGCCGA
35	12721	TCAGCGTGAG	CGTGGCGGTG	TCGGACCGGG	CCAGCAGCCG	GCTCGCGACG	CGCGCGACCG
	12781	CGCCGCTCCG	CATCGCGGTG	ATCACGCCCTG	CGTCGCGCAG	GGCGGTGAGA	CTGCCGCTGT
	12841	CGTCGTGAG	GCGCGACATC	GTGCCGACGA	TCGTGCGCAG	CCGGAAGCGC	GGATAGTTGT
	12901	GCGGACTGTA	CGAAACCGTC	TTCATGGTCA	CGCCGACACC	GGGACCCCGG	TACGGCATGA
	12961	ACTCGATGAC	GCCGGGAATG	TGCGCGCCGC	GGACGAATCC	GGTACGCGGC	GGCGCCTCGG
40	13021	CGAACTCGCC	GCGGCCGAGC	GCGGCGAACC	CGTCGTGACG	CTCGCTGATC	AGCCGGTCCA
	13081	TCATCAGCTC	GCGGCCGATC	ACGGAGAGAA	TCCGCTTGAT	GTCACGTTGG	CGCAGGACCC
	13141	TGGTCTGCAT	GTGTCACCTC	CCTTTCTGTTG	CCGGAGCTGT	CTTGGTGGTG	CGGCTCGGGG
	13201	CGGCTTCCGT	TCTCATCGCA	GCTCCCTGTC	GATGAGGTG	AAAATCTCGT	CCGCGGTGCG
	13261	GTCCGCGGAC	AGCACGCCGG	CCGGCGTGGT	CGGGCGGGTC	TCCCGCCGCC	AGCGGTTGAG
45	13321	CAGGGCGTCC	AGCCGGGTTT	CGATCGCGTC	CGCCTGGCGG	GCGCCCGGGT	CGACACCGGC
	13381	AACGAGTGCT	TCCAGCCGGT	CGAGCTGCGC	GAGCACCACG	GTCACCGGGT	CGTCCGGGGA
	13441	CAGCAGTTCA	CCGATGCGGT	CGGCGAGTGC	GCGCGGCGAC	GGGTAGTCGA	AGACGAGCGT
	13501	GGCGGACAGT	CGCAGACCGG	TCGCCTCGTT	GAGGCCGTTG	CGCAGCTGCA	CCGCGATGAG
	13561	CGAGTCCACA	CCGAGTTCCC	GGAACGCGCG	GTCTCTCCGG	ATGTCCTCCG	GGTCGCGGTG
50	13621	GCCCAAGACG	GCCGCTGCCT	TCTGCCGGAC	GAGGGCGAGC	AGGTCCGGTG	GGCGTCTCTG
	13681	CTCGTTGCGG	GCGCTCCGGC	GGGCCGACGG	CTTGGGCCGG	CCACGCAGCA	GCGGGAGGTC
	13741	CGGCGGCGAG	TCGCCCAGCA	CGGCGACGAC	ACTGCCCGTT	CCGGTGTGGA	CGGCGCGGTC
	13801	GTACATGCGC	ATGCCCTGTT	CGGCGGTGAG	CGCGCTCGCC	CCACCCTTGC	GCATACGGCG
	13861	CCGGTCCGGC	TCGGTCAGGT	CCGCGGTGAG	GCCACTCGCC	TGGTCCCACA	GCCCCACGCG
55	13921	GATCGACAGC	CCTGGCAGCC	CTTGTGCACG	CCGGTGTTCG	GCGAGCGAGT	CCGAGAACGC
	13981	GTTCCGCGCC	GCGTAGTTGC	CCTGACCGGG	GGTGCCGAGC	ACACCGGCCG	CCGACGAGTA
	14041	GACGACGAAT	GCGGCGAGGT	CGGTGTGCGG	GGTGAGCCGG	TGCAGGTGCC	AGGCGGCGTC
	14101	GGCCTTGGGT	TTGAGGACGG	TGTCGATGCG	GTCGGGGGTG	AGGTTGTGCA	GCAGGGCGTC
	14161	GTCGAGGGTT	CCGGCGGTGT	GGAAGACGGC	GGTGAGGGGT	TGAGGGATGT	GGGCGAGGGT
60	14221	GGTGGCGAGT	TGGTGGGGGT	CGCCGACGTC	GCAGGGGAGG	TGGGTGCCGG	GGGTGGTGTC
	14281	GGGGGGTGGG	GTGCGGGAGA	GGAGGTAGGT	GTGGGGGTGG	TTCAGGTGGC	GGGCGAGGAT
	14341	CGCGGCGAGG	GTGCGGAGC	CGCGGTGAT	GACGACGGCC	CCCTCGGGGT	CCAGCGGCGG
	14401	CGGGACCGTG	AGGACGATCT	TGCCGGTGTG	CTCGCCGCGG	CTCATGGTCC	CCAGCGCCTC
	14461	GCGGACCTGC	CGCATGTCTG	GCACCGTCAC	CGGCGAGGGG	TGCAGCACAC	CGCGCGCGAA

	14521	CAGGCCGAGC	AGCTCCGCGA	TGATCTCCTT	GAGCCGGTGC	GGCCCCGCGT	CCATCAGGTC
	14581	GAACGGTTCG	TGGACGGCGT	GCCGGATGTC	CGTCTTCCCC	ATCTCGATGA	ACCGGCCACC
	14641	CGGCGCGAGC	AGGCCGACGG	ACGCGTCGAG	GAGTTCACCG	GTGAGCGAGT	TGAGCACGAC
5	14701	GTCCAGCGGC	GGGAACGCGT	CGGCGAACGC	GGTGCTGCGG	GAATCGECCA	GATGCGCTCC
	14761	GTCCAGGTCC	ACCAGATGGC	GCTTCGCGGC	GCTGGTGGTC	GCGTACACCT	CCGCGCCCAG
	14821	GTGCCGCGCG	ATCTGCCGGG	CGGCGGAACC	GACACCGCCG	GTGGCCGCGT	GGATCAGGAC
	14881	CTTCTCGCCG	GGGCGCAGCC	CGGCGAGGTC	GACCAGGCCG	TACCACGCGG	TCGCGAACGC
	14941	GGTCATCACG	GACGCCGCCT	GCGGGAACGT	CCAGCCGTCC	GGCATCCGGC	CGAGCATCCG
10	15001	GTGGTCGGCG	ATGACCGTGG	GGCCGAAGCC	GGTGCCGACG	AGGCCGAAGA	CGCGCTGGCC
	15061	CGGTGCCAGA	CCGGAGACGT	CGGCGCCGGT	CTCCAGGACG	ATGCCCGCGG	CCTCGCCGCC
	15121	GAGCACGCCC	TGACCGGGGT	AGGTGCCGAG	CGCGATCAGC	ACATCGCGGA	AGTTGAGGCC
	15181	CGCCGCGACG	ACACCGATCC	GGACCTCGGC	CGGGGCGAGG	GGGCGCCGGG	GCTCCGCCGA
	15241	GTCGGCCGCG	GTGAGGCGGT	CGAGGCTGCC	CGTCCGCGCC	GGCCGGATCA	GCCACGTGTC
	15301	GCTGTCCGGC	ACGGTGAGCG	GCTCCGGCAC	CCGGGTGAGG	CGGGCCGCCCT	CGAACCGGCC
15	15361	GCCGCGCAGC	CGCAGACGCG	TGCGGCGGAG	TCGCGACGGC	ATGCGCTGCT	GCTCGGGGGC
	15421	GAGCGTGACG	CCGGACTCGG	TCTCGACGTG	GACGAACCGG	CCGGGCTGCT	CCGGCTGGGC
	15481	GGCGCGCAGC	AGTCCGGCCG	CCGCGCCGGT	GGCGAGGCCG	GCGGTGGTGT	GCACGAGCAG
	15541	ATCCCCGCCG	GAGCCGGTCA	GGGCGGTGAG	CAGCCGGGTG	GTGAGCGCAC	GCGTCTCGGC
20	15601	CACCGGGTCG	TCGCCATCAG	CGGCAAGCAA	CGTGATGACG	TCCACGTCCG	TCGCGGGGAC
	15661	ATCCCTGGGT	GCGGCGACCT	CGATCCAGGT	GAGACGCATC	AGGCCGGTGC	CGACGGGTGG
	15721	GGACAGCGGG	CGGGTGCGGA	CCGTCCGGAT	CTCGGCGACG	AGTTGGCCGG	CGGAGTCGGC
	15781	GACGCGCAGA	CTCAGTCTGT	CGCCCTCAGC	AGTGATCAGC	GCTCGGAGCA	TGGCCGAGCC
	15841	CGTGCGGACG	AACCGGGCCC	CCTTCCAGGC	GAACGGCAGA	CCCGCAGCGC	TGTCGTCCGG
25	15901	CGTGGTGAGG	GCGACGGCGT	GCAGGGCCGC	GTCGAGCAGC	GCCGGATGCA	CACCGAAACC
	15961	GTCCGCCCTCG	GCGGCTGTGT	CGTCCGGCAG	CGCCACCTCG	GCATACACGG	TGTCACCATC
	16021	ACGCCAGGCA	GCCCCGAACC	CCTGGAACGC	CGACCCGTAC	TCATAACCGG	CATCCCGCAG
	16081	TTCGTTCATAG	AACCCCGAGA	CGTCGACGGC	CACGGCCGTG	ACCGGCGGCC	ACTGCGAGAA
	16141	CGGCTCCACA	CCGACAACAC	CGGGGTGTGT	GGGGGTGTGT	GGGGTCAGGG	TGCCGCTGGC
30	16201	GTCGCCGGTC	CAGCTGCCCG	TGCCCTCGGT	ACGCGCGTGG	ACGGTCACCG	GCCGCCGTCC
	16261	GGCCTCATCA	GCCCCCTTCCA	CGGTACCCGA	CACATCCACC	GCTGCGGTCA	CCGGCACCC7.C
	16321	AAGGGGGGAT	TCGATGACCA	GCTCGTCCAC	TATCCCGCAA	CCGGTCTCGT	CACCGGCCCG
	16381	GATGACCAGC	TCCACAAACG	CCGTACCCGG	CAGCAGGACC	GTGCCCCGCA	CCGCGTGATC
	16441	AGCCAGCCAG	GGGTGAGTGC	GCAATGAGAT	CCGGCCAGTG	AGAACAACAC	CACCATCGTC
35	16501	GGCGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CGGATGCGCC	GCACCCGTCA	ACCCCGCCGC
	16561	CGACAGATCG	GTGGACCCGG	CCGCTCCAG	CCAGTACCGC	CTGTGCTCGA	ACCGGTACGT
	16621	GGGCGATCC	AGCAGCCGTC	CCGGCACCAG	TTCGACCACC	GTGTCCAGT	CCACTGCCGT
	16681	GCCAGGGTTC	CACGCCCTGCG	CCAACGCCGT	CAGCCACCGC	TCCAGCCGCG	CGTCACCGGT
	16741	CCGCAACGAC	GCCACCGTGT	GAGCCTGCTC	CATCGCCGGC	AGCAGCACCG	GATGGGCACT
40	16801	GCACTCCACG	AACACCGACC	CATCCAGCTC	CGCCACCGCC	GCGTCCAACG	CCACCGGACG
	16861	ACGCGAGATC	CGGTACCACT	ACCCCTCATC	CACCGGCTCC	GTCACCCAGG	CGCTGTCCAC
	16921	GGTCGACCAC	CACGCCACCG	ACGCGGCCCT	CCCTGCCACC	CCCTCCAGTA	CGTTGGCCAG
	16981	TTATCCTCTG	ATGGCTTCCA	CGTGGGGCGT	GTGGGAGGCG	TAGTCAGCCG	CATACGACG
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
45	17101	CGCCACCACC	GTCGAAGCCG	GGCCGTTACG	CGCCGCGATC	CACACACCCT	CGACCAGACC
	17161	GACCTCACC	GCCGGCAACG	CCACCGAAGC	CATCGCTCCC	CGCCCGGCCA	GTCGCGCCGC
	17221	GATGACCTGA	CTGCGCAATG	CCACCACGCG	GGCGGCGTCC	TCGAGGCTGA	GGGCTCCGGC
	17281	CACGCACGCC	GCCGCGATCT	CGCCCTGGGA	GTGTCCGATC	ACCGCGTCCG	GCACGACCCC
	17341	ATGCGCCTGC	CACAGCGCGG	CCAGGCTCAC	CGCGACCGCC	CAGTGGCCCG	GCTGGACCC
50	17401	CTCCACCCGC	TCCGCCACAT	CCGGCCGCGC	CAACATCTCC	CGCACATCCC	AGCCCCGTGT
	17461	CGGCGAGCAAC	GCCTGAGCGC	ACTCCTCCAT	ACGCGCGGCG	AACACCGCGG	AGTGGGUCAT
	17521	GAGTTCACG	CCCATGCCGA	CCCACTGGGC	GCCCTGGCCG	GGGAAGACGA	ACACCGTACG
	17581	CGCCTGGTCC	ACCGCCACAC	CCGTACCCCG	GGCATCGCCC	AGCAGCACCG	CACGGTGACC
	17641	GAAGACAGCA	CGCTCCCGCA	CCAACCCCTG	CGCGACCGCG	GCCACATCCA	CACCACCCCC
55	17701	GCGCAGATAC	CCCTCCAGCC	GCTCCACCTG	CCCCCGCAGA	CTCACCTCAC	CACGAGCCGA
	17761	CACCGGCAAC	GGCACCACAC	CGTCAACAAC	CGACTCCCCA	CGGACGCGCC	CAGGAACACC
	17821	CTCAAGGATC	ACGTGCGCGT	TGCTACCGCT	CACCCCGAAC	GACGACACAC	CCGCATGCGG
	17881	TGCCCCGATCC	GACTCGGGCC	ACGGCCTCGC	CTCGGTGAGC	AGCTCCACCG	CACCGGCCGA
	17941	CCAGTCCACA	TGCGACGACG	GCTCGTCCAC	ATGCAGCGTC	TTCGGCGCGA	TCCCGTACCG
60	18001	CATCGCCATG	ACCATCTTGA	TCACACCGGC	GACACCCGCC	GCCGCTGCG	CATGACCGAT
	18061	GTTCGACTTC	AACGAACCCA	GCAGCAGCGG	AACCTCACGC	TCCTGCCCGT	ACGTCGCCAG
	18121	AATGGCCTGC	GCCTCGATGG	GATCGCCAG	CGTCGTCCCC	GTCCCGTGCG	CCTCCACCAC
	18181	GTCACATCG	GCGGCGCGCA	GTCGCGCGTT	CACCAACGCC	TGCTGGATGA	CACGCTGCTG
	18241	GGACGGGCGG	TTGGGGGCGG	ACAGCCCGTT	GGAGGCACCG	TCCTGGTTCA	CCGCCGACCC
	18301	GCGGACGACC	GCGAGAACGG	TGTGTCCGTT	GCGCTCGGCG	TCGGAGAGCC	GCTCCAGCAC

	18361	AAGAACGCCG	GCGCCCTCCG	CCCAGCCGGT	GCCGTGGGCG	GCGTCCGCGA	ACGCGCGGCA
	18421	GCGGCCGTCTG	GGGGAGAGTC	CGCCCTGCTG	CTGGAATTCC	ACGAACCCGG	TCGGGGTCCG
	18481	CATGACGGTG	ACACCGCCGA	CCAGCGCCAG	CGAGCACTCC	CCGTGGCGCA	GTGCGTGGCC
	18541	GGCCTGGTGC	AGCGCGACCA	GCGACGACGA	GCACGCCGTG	TCCACCGTGA	ACGCCGGTCC
5	18601	CTGGAGCCCA	TAGAAGTACG	AGATCCGGCC	GGTGAGCACG	CTGGGCTGCA	TGCCGATCGA
	18661	GCCGAACCCG	TCCAGGTCCG	CGCCGACGCC	GTACCCGTAC	GAGAAGGCGC	CCATGAACAC
	18721	GCCGGTGTCTG	CTGCCGCGCA	GTGTGCCCGG	CACGATGCCC	GCGCTCTCGA	ACGCCTCCCA
	18781	TGTCGTTTTCC	AGCAGGATCC	GCTGCTGGGG	GTCCATGGCC	CGTGCCCTCAC	GGGGGCTGAT
	18841	GCCGAAGAAC	GCGGCATCGA	AGCCGGCCGG	CTCGGAGAGG	AAGCCGCCCG	GGTCCGTGTC
10	18901	CGATCCGCCG	GTGAGGCCGG	ACGGGTCCCA	GCCACGGTCG	GCCGGGAAGC	CGGTGACCGC
	18961	GTCGCCGCCA	CTGTCCACCA	TGCGCCACAG	GTCGTGCGGC	GAGGTGACGC	CGCCCGGCAG
	19021	TCGGCAGGCC	ATGCCCACGA	TGGCCAGCGG	TTCGTACCGG	GTCGCGCGCG	CTGTGGGAAC
	19081	AGCGACCGGT	GCGGCACCAC	CGACCAGAGC	CTCGTCCAAC	CGCGACGCGA	TGGCCCGCGG
	19141	CGTCGGGTAG	TCAAGACAA	GCGTGGCGGG	CAGTCGGACA	CCGGTCGCCG	CGGCGAGTCG
15	19201	GTTCCGCAGT	TCGACGGCGG	TCAGCGAGTC	GATACCCAGT	TCCTTGAAGG	CCGCGTCCCG
	19261	GCGACGTCC	GCGGCGTCCG	CGTGGCCGAG	CACCGCCGCC	GCGTTGTCTG	GGACCACTGC
	19321	CAGCAGCGCG	GTGTCCCGCT	CAGCGCCGGA	CATGGTGCCG	AGCCGGTCGG	CGAGCGGAAC
	19381	GGCGGTGGCC	GCCGCCGGGC	GCGATACGGC	GCGGCGCAGA	TCGGCGAAAA	GCGGCGATGT
	19441	GTGCGCGGTG	AGGTCCATCG	TGGCCGCCAC	GGCGAACGCG	GTGCCGGTTC	CGGCCGCGGC
20	19501	TTCCAGCAGG	CGCATGCCCA	CACCGGCCGA	CATGGGGCGG	AAACCGCCCG	GGCGGACACG
	19561	GGTGCGGTTG	GTGCCGCTCA	TGCTGCCGGT	GAGTCCGCTG	TCATCGGCCC	AGAGGCCCCA
	19621	GGCCAGCGAC	AGCGCGGGCA	GTCTTCCGGC	ATGGCGCAGC	GTCGCGAGTC	CGTCGAGGAA
	19681	CCCGTTCCGC	CCCGAGTAGT	GTCCTGGGCC	GCGGCGCGCC	ATGATGCGCG	CGAGAGCGCA
	19741	GTAGAGGACG	AACGAGCGCA	GGTCCGCGTC	CCGGGTCAGC	TCGTGCAGGT	GCCAGGCGCC
25	19801	GTCGGCTTTG	GGGCGCAGTG	TGGTGGCGAG	CCGCTCCGGG	GTGAGTGCCG	TGGTCACGCC
	19861	GTCGTGAGC	ACGGCTGCCG	TGTGGAAGAC	CGCCGTGAGC	GGCCTGCCGG	CGGCGGCGAG
	19921	CGCGGCGGGC	AGCTGGTCCC	GGTCGGCGAC	GTCACAGCGG	ATGTGGACAC	CGGGAGTGTC
	19981	CGCCGGCGGT	TCGCTGCGCG	ACAGCAACAG	GAGGTGGCGG	GCGCCATGCT	CGGCGACGAG
	20041	ATGCGGGCGG	AGGAGACCTG	CCAGCAACCC	CGAGCCGCCG	GTGATGACCA	CCGTGCCGTC
30	20101	CGGGTCGAGC	AGCGGTTCCG	GCGTTTCCGC	GGCGGCCGTG	CGGGTGAACC	GCGGCGCTTC
	20161	GTACCGGCCG	TCGGTGACGC	GGACGTACGG	CTCGGCCAGT	GTCGTGGCGG	CGGCCAGCCC
	20221	CTCGATGGGG	GTGTGCGTGC	CGGTCTCCAC	CAGCACGAAC	CGGCCCGGGT	GCTCGGCTTG
	20281	GGCGGACCGG	ACGAGGCCGG	CGACCGCTCC	TCCGACCGGT	CCCGCGTCGA	TCCGACGAC
	20341	GAGGGTGGTC	TCCGCAGGGC	CGTCTCCGGC	GATACCCCGG	TGCAGCTCGC	CGAGCACGAA
35	20401	CTCGGTGAGC	CGGTACGTCT	CGTCGAGGAC	ATCCGCGCCC	GGTTCGGGGA	GCGCGGAGAC
	20461	GATCTGGACC	GCGTCCGACT	GACCGGGCCC	GGGAGTGGGC	AGTCCGGTCC	AGGAGAGGCC
	20521	GTACAAGGAG	TTCCGTACGA	CGGCGGCGTC	GCCGTGACG	TTCACCGGTC	GCGCGGTGAG
	20581	CGCGGCGACG	GTCACCACCG	GTTGGCCGAC	CGGGTCCGTC	GCATGCACGG	CAGCGCCGTC
	20641	CGGGCCCTGA	GTGATCGTGA	CGCGCAGCGT	GGTGGCCCCG	GTCGTGTGGA	ACCGCACGCC
40	20701	GCTCCACGAG	AACGGCAGCC	GCACCTCCGC	TTCCTGTTCC	GCGAGCAGCG	GCAAGCAGGT
	20761	GACGTGCAAG	GCCGCGTCTGA	ACAGCGCCGG	GTGGACGCCA	TAGTGCGGCG	TGTCGTCEGC
	20821	CTGTTCCCCG	GCGATCTCCA	CCTCGCGGTA	CAGGGTTTCG	CCGTGCGCGC	AGGCGGTGCG
	20881	GAGTCCCTGG	AACGTGCGGC	CTAGCTGTA	GCCGGTCTCG	GCCAGCCGCT	CGTAGAACCG
	20941	GCTCACGTCTG	ACGCGTCTCG	CGCCCGCGCG	CGGCCACGCG	GGCGGCGGGA	CCGCCGCGAC
45	21001	GCTTCCGGCC	CGGCCGAGGG	TGCCGCTGGC	GTGCCGGGTC	CAGCTGTCCG	TGCCCTCGGT
	21061	ACGCGCGTGG	ACGGTCACTC	GCCGCCGTCC	GGCCTCATCG	GCCCCTTCGA	CGGTACCCGA
	21121	CACATCCACC	GCGCCGGTCA	CCGGCACACC	GAGCGGGGTC	TGCATGACCA	GTTATCCACC
	21181	CACCCCGCAA	CCGGTCTCGT	CACCGGCCCG	GATGACCAGC	TCCACAAACG	CCGTACCCGG
	21241	CAGCAGAACC	GTGCCCGCA	CCCGTGATC	AGCCAGCCAG	GGATGCGTAC	GCAACGAGAT
50	21301	CCGGCCAGTG	AGAACAACAC	CACCACCGTC	GTCGGCGGGC	AGTGCTGTGA	CGGCGGCCAG
	21361	CATCGGATGC	GCCGCCCCCG	TCAGCCCGGC	CGCGGACAGA	TCCGTGGCAC	CGGCCGCTTC
	21421	CAGCCAGTAC	CGCCTGTGCT	CGAACGCGTA	GGTGGGCGAG	TCCAGCAGCC	GTCCCGGCAC
	21481	CGGTTTCGACC	ACCGTGTCCC	AGTCCACTGC	CGTGCCCGAG	GTCCACGCCT	GCGCCAACGC
	21541	CGTCAGCCAC	CGCTCCCAGC	CGCCGTCAAC	GGTCCGCAAC	GACGCCACCG	TGTGAGCCTG
55	21601	TTCCATCGCC	GGCAGCAGCA	CCGGATGGGC	GCTGCACTCC	ACGAACACGG	ACCCGTCCAG
	21661	CTCCGCCACC	GCCGCGTCCA	GCGGACGAGG	GCGACGACGG	TTCCGGTACC	AGTAGCCCTC
	21721	ATCCACCGGC	TCGGTCAACC	AGGCGCTGTC	CACCGTGGAC	CACCAGGCCA	CCGACCCGGT
	21781	CCCGCCGGAA	ATCCCTCCA	GTACCTCGGC	CAACTCGTCC	TCGATGGCTT	CCACGTGGGG
	21841	CGTGTGGGAG	GCGTAGTCTGA	CCGCGATACG	GCGCACTCGC	ACGCCTTCGG	CCTCGTACCG
60	21901	CGTCACCACT	TCTTCCACCG	CGGACGGGTC	CCCCGCCACC	ACAGTCGAAG	ACGGGCCGTT
	21961	ACGCGCCGCG	ATCCACACGC	CCTCGACCAG	GTCCACCTCA	CCGGCCGGCA	ACGCCACCGA
	22021	AGCCATCGCC	CCCCGCCCGC	CCAGCGCCCG	GCGGATCACC	TGGCTGCGCA	AGGCCACCGA
	22081	GCGGGCGGCG	TCCTCAAGGC	TGAGGGCTCC	GGCCACACAC	GCCGCGCGCA	TCTCGCCCTG
	22141	GGAGTGTCCG	ACCACCGCGT	CCGGCACGAC	CCCATGCGCC	TGCCACAGCG	CGGCCAGGCT

	22201	CACCGCGACC	GCCCAGCTGG	CCGGCTGGAC	CACCTCCACC	CGTCCGCCA	CATCCGGCCG
	22261	CGCCAACATC	TCCCGCACAT	CCCAGCCCGT	GTGCGGCAAC	AACGCCCCG	CACACTCCTC
	22321	CATACGAGCC	GCGAACACCG	CAGAACACGC	CATCAACTCC	ACACCCATGC	CCACCCACTG
5	22381	AGCACCTGTC	CCGGGAAAGA	CGAACACCGT	ACGCGGCTGA	TCCACCGCCA	CACCCATCAC
	22441	CCGGGCATCG	CCCAACAACA	CCGCACGGTG	ACCGAAGACA	GCACGCTCAC	GCACCAACCC
	22501	CTGCGCGACC	GCGGCCACAT	CCACACCACC	CCCGCGCAGA	TACCCCTCCA	GCCGCTCCAC
	22561	CTGCCCCCGC	AGACTCACCT	CACCTCCGAGC	CGACACCGGC	AACGGCACCA	ACCCATCGAC
	22621	AGCCGACTCC	CCACGCGACG	GCCCAGGAAAC	ACCCTCAAGG	ATCACGTGCG	CGTTCGTACC
10	22681	GCTCACCCCG	AAAGCGGAGA	CACCGGCCCG	GCGCGGACGT	CCCGCGTCGG	GCCACGCCCG
	22741	CGCCTCGGTG	AGCAGTTCCT	CCGCGCCCTC	GGTCCAGTCC	ACATGCGACG	ACGGCTCGTC
	22801	CACATGCAGC	GTCTTCGGCG	CGATGCCATA	CCGCATCGCC	ATGACCATCT	TGATGACACC
	22861	GGGACACCC	GCAGCCGCCT	GCGCATGACC	GATGTTTCGAC	TTCAACGAAC	CCAGCAGCAG
	22921	CGGAACCTCA	CGCTCCTGCC	CGTACGTCGC	CAGAATCGCG	TGCGCCTCGA	TGGGATCGCC
15	22981	CAGCGTCGTC	CCCGTCCCGT	GCGCCTCCAC	ACGCTCCACG	TGCGCGGGGG	CGGACCCJCGC
	23041	CTTGTGGAGG	GCCTGGCGGA	TGACGCGCTG	CTGGGAGGGG	CCGTTGGGTG	CGGAGATGCC
	23101	GTTGAGGCG	CCGTCTTGGT	TGACGGCGGA	GGAGCGGACG	ACCGCGAGGA	CGGTGTGTCC
	23161	GTTGCGCTCG	GCGTCGGAGA	GCTTTTCGAC	GACGAGGACG	CCGGCCCCCT	CGGCGAAACC
	23221	GGTGCCGTCC	GCCGCGTCAG	CGAACGCCTT	GCACCGTCCG	TCCGGCGCGA	CGCCGCCCTG
20	23281	CCGGGAGAAC	TCCACGAAGG	TCTGTGGTGA	TGCCATCACT	GTGACACCAC	CGACCAGCGC
	23341	CAGCGAGCAC	TCCCGGTCC	CGACGCGCTG	CCCGGCTTGG	TGCAGCCGCA	CCAGCGACGA
	23401	CGAACACGCC	GTGTCGACCG	TGACCGCCGG	ACCCTCCATG	CCGAAGAAGT	ACGACAGCCG
	23461	TCCGGCGAGC	ACCGCGGGCT	GTGTGCTGTA	GGCGCCGAAT	CCGCCAGGT	CCGCGCCCGT
	23521	GCCGTAGCCG	TAGTAGAAGC	CGCCGACGAA	GACGCCGGTG	TCGCTGCCGC	GCAGGGTGTC
25	23581	CGGCACGATG	CCGGCGTGTT	CGAGCGCCTC	CCAGGCGATT	TCGAGGAGGA	TCCGCTGCTG
	23641	CGGGTCGAGT	GCGGTGGCCT	CGCGCGGACT	GATGCCGAAG	AACGCGGCAT	CGAAGTCGGC
	23701	GGCGCCCGCG	AGTGCGCCGG	CCCGCCCGGT	GGCGGACTCG	GCGGCGGCGT	GCAGCGCGGC
	23761	CACGTCCCAG	CCGCGGTCCG	TGGGGAAGTC	GCCGATCGCG	TGCGGCGCGT	CCGCGACGAG
	23821	CTGCCACAGC	TCTTCCGGTG	AGGTGACGCC	GCCCCGCAGT	CGGCAGGCCA	TGCCGACGAC
30	23881	GGCGAGCGGC	TCGTTCCCGG	CGGCGCGCAG	CGCGGTGTTC	TCCCGGCGGA	GCTGCGCGTT
	23941	GTCTTTCGAC	GACGTCCGCA	GCGCCTCGAT	CAGGTGCTTC	TGCGCCATCG	CCTCATCCCT
	24001	TCAGCACGTG	CGCGATGAGC	GCGTCTGCGT	CCATGTCTGTC	GAACAGTTTC	TGCTCCGGCT
	24061	CCGCGGTCTG	GGTGCTCGCG	GGTGCCTGTG	CCGGTGGTTC	ACCGCCGTCC	GGGGTCCCGT
	24121	TGTGCTCCGG	GGTCCCCTTG	ACGTCCGGGG	CCAGGAGGGT	CAGCAGATGA	CGGGTGAGCG
35	24181	CGCCGGCGGC	GGGATAGTCG	AAGACGACGC	TGGCCGGCAG	CGGAATGCCG	AGGGCTCCGG
	24241	AGACCCGGTT	GCGCAGGCCG	AGCGCGGTGA	GCGAGTCGAC	CCCGAGGTCC	TTGAACGCCG
	24301	TGGTGGCCGT	GACCGCCGCC	GCGTCGGTGT	GGCCAGCAG	GGTGGCGGCG	GTGTGCGGGA
	24361	CGACGCCGAG	CAGCACCTGT	TCCCGTTCCT	TGTGGGGCAG	GTCCGGCAGG	CGTTCAGCA
	24421	GGGAGCCGCC	GTCGGTCCGG	GAGCGCCGGG	TGGGGCGCTG	GATCGGTTCG	CACAGCGGTG
40	24481	ACGGGTCCGC	GGGCCCCGGT	GGGGCGGTTC	CCACGACCAC	GGCTTCCCCG	GTGGCGCACG
	24541	CGGCGTCGAG	GAGGTCCGTC	AGCCGGTCCG	CCGCGGCGGT	GAACGCCACG	GCCGCGAGGC
	24601	CTTGTGCCCC	GCGCAGGTTC	GCCAGGGCCT	GGAGCGGTCC	GGCCGCTTCG	CCGGACGGAA
	24661	CGGCGAGAAC	GAACGCGGTC	AGGTGAGAGT	CGCGGGTCAG	GCGGTGCAGT	TCCCAGGCCG
	24721	ACTCGGCGGT	GCCGTCCGCG	TGGACGACCG	CGGTCAACGG	GGTTTCCGGC	ACTGTGCCCG
45	24781	GCTCGTACCG	GATCACTTCG	GCGCCGTGTC	CGCCGAGGTG	TCCGGCGAGT	TCCTCCGAAC
	24841	CGCCCGCGAG	GAGGACGGTG	TCGCCGTACG	AGGCCGCGGC	CGTGGTGGGC	GCGGCGGGGA
	24901	CGAGGCGGGG	CGCTTCGAGG	CGCCCGTCCG	CCAGGCGCAG	GTGCGGTTCG	TCGAGGCGGG
	24961	AGAGGCGGGC	GCGCGGCGCG	GGGGTGACCG	TGTCGGTGGT	CTCCACGAGC	ACGAGCCGGC
	25021	CCGTTTCCGC	GGTGTGAGC	AGTGCGGCGA	CGGCACCGGC	GACGGGCCCC	GCCCTCGGCG
50	25081	ACACCACCAG	CGTGGCGCCG	GCGGTCTCTG	GGTCTGTCAG	TGCGGTACGG	ACCTCGTCGG
	25141	GACCGGATAC	CGGGACGACG	ATGACGTCCG	GCGTGGCGTC	GTGCGCGAGG	TCGGTGTACC
	25201	GGCGGGCCGT	GGTGCCGGGT	GCCGCGGGGG	CCCGGACGCC	GGTCCAGGTG	CGCCGGAACA
	25261	GCCGCACGTC	CCCGTCCGGG	CCCGTCCGTT	CGGGGGGCGG	GGTGATGAGC	GAGCCGCTCT
	25321	GAGCCACCGG	CCGTCCAGT	TCGTCCGCGA	GGTGACGCG	GGCGCCGCC	TGCCCTTCGC
55	25381	GCTGGACGAA	GGTGACGCGC	AGTTTCGTGG	GCGCGCTGGT	GTGGACACGG	ACGCGGTGA
	25441	ACGCGAACGG	CAACCGTACC	CCCGCGTTCT	CGGCGGCCGC	GCCGATGCTG	CCCGCTTGCA
	25501	GCGCGGTGAC	GAGCAGCGCC	GGGTGCAGTG	TGTAGCGGGC	GGCGTCCCTG	GCGAGGGCGC
	25561	CGTCGAGGGC	GACTTCGGCG	CAGACGGTGT	CTCCGTGGCT	CCACGCGGGC	GACATGCCGC
	25621	GGAACTCGGG	GCCGAACCTC	TATCCCGCGT	CGTCGAGTCG	CTGGTAGAAG	GCCGCGACGT
60	25681	CGACCGGTTT	CGCGTGCTCG	GGCGGCCAGG	GCCCCGGCGT	GGTGGCCGGT	TCGGTGGTGG
	25741	CGATGCCGGC	GAAGCCGGAG	GCGTGCGGGG	TCCATGTCCG	GTGCGCGTCC	GTCCGGGCGT
	25801	GACGCGGCAC	GGCACGGCGT	CCGGTGTCTG	CGGGCGCGGC	GACGGTCACG	CGACCTGGA
	25861	CGGCGCCGGT	GGCGGGCAGG	ACCAGCGGTG	TCTCGACGAC	CAGTTCGTCT	AGCAGGTCTG
	25921	AGCCTGCCTC	GTCGGCGCCG	CGTCCGGCCA	ATTCCAGGAA	GGCGGGTCCG	GGCAGCAGTA
	25981	CGGCGCCGTC	GACGGAGTGA	CCGGCCAGCC	ATGGGTGGGT	GGCCAGCGAG	AACCGGCCCG

26041 TGAGCAGCAC CTCGTCGGAG TCGGGGAGCG CCACCGACGC GCGAGCAGC GGGTGGTCTGA  
 26101 CGGCGTCGAG TCCGAGGCCG GAAGCGTCCG TGCCGGCCGC GGTCTCGATC CAGTAGCGCT  
 26161 CATGGTGGAA GGCGTATGTG GGCAGGTCGT GTGCCGTGCG CGTCGCGGGG ACGACCGCCG  
 26221 CCCAGTCGAC GGGCACGCCG GTTGTGTGCG CCTCGGCCAG CGCGGTGAGC AGCCGGTGGGA  
 5 26281 CTCCCCCGCC GCGGCGGAGC GTGGCGACGG TCGCGCCGTC GATCGCGGGC AGCAGCACGG  
 26341 GGTGCGCGCT GACCTCGACG AACACGGTGT CACCCGGCTC GCGGGCAGCG GTCACGGCCG  
 26401 TGGCGAAGCC TACGGGGTGG CGCATGTTGC GGAACAGTA CTCGTCGTG AGCGGCGCGT  
 26461 CGATCCAGCG TTCGTCGGCG GTGGAGAACC ACGGGATCTC GGGCGTGCGG GAGGTGGTGT  
 10 26521 CCGCGACGAT CCGCTGGAGT TCGTCGTACA GCGGGTCGAC GAACGGGGTG TGGTCCGGG  
 26581 AGTCGACGGC GATGCGGCGC ACCCAGACGC CGCGGGCCTC GTAGTCGGCG ATCAGCGTTT  
 26641 CGACGGCGTC CGGGCGCCCG GCGACGGTCG TGGTGGTGGC GCCGTTGCGG CCCGCGACCC  
 26701 AGACGCCGTC GATCCGGGCG GCATCCGCTC CGACGTCGGC GGCCGGGAGC GCGACCGAGC  
 26761 CCATCGCGCC GCGTCCGGCG AGTTCGCGCA GGAGCAGGAG AACGCTGCGC AGCGCGACGA  
 26821 GGGGGGACCC GTCCTCCAGG GTGAGCGCTC CGGCGACACA GGCCGCGGGC ATCTCGCCCT  
 15 26881 GGGAGTGTCC GATGACGGG TCCGAGCGTA CGCCCGCGGC CTCCCACACG CGGGCCAGCG  
 26941 ACACCATGAC GGCCAGCAG ACGGGGTGCA CGACGTCGAC GCGGCGGGT ACCTCCGGGT  
 27001 CGTCGAGCAT GGCGATGGGG TCCAGCCCCG TGTGCGGGAT CAGCGCGTCG GCGCATTGGC  
 27061 GCATCCTGGC GGCGAACACC GGGGAGGCCG CCATCAGTTC GACGCCCATG CCGCGCCACT  
 27121 GCGGTCTTTG TCCGGGGAAG ACGAAGACGG TGCGCGGCTC GGTGAGCGCC GTGCCGGTGA  
 20 27181 CGACGTCGTC CTCGAGCAGC ACGGCGCGGT GCGGGAACGT CGTACGCCTG GCGAGCAGGC  
 27241 CCGCGCGCAT GCGCGCGGGT TCGTGGCCGG GACGGGCGGC GAGGTGCTCG CGGAGTCGGC  
 27301 GGACCTGGCC GTCGAGGGCC GTGGCGGTCG CGCCCGAGAC GGGCAGTGGT GTGAGCGGCG  
 27361 TGGCGATCAG CCGCTACCG GGCTTCGAGG CCGACGGCTC CTCGGCCGGC GGCTCCCGCG  
 27421 CCGGGTGGGC TTCCAGCAGG ACGTGGGCGT TGGTGCCGCT GACGCCGAAG GAGGACACAC  
 25 27481 CCGCGCGCCG CCGGCGGTCG GTCTCGGGCC AGGGCCGGGC ATCGGTGAGG AGTTCGACGG  
 27541 CGCCGGCCGT CCAGTCGACG TGCGAGGACG GCGTGTCCAC GTGCAGGGTG CGCGGCAGGG  
 27601 TGCCGTGCCG CATGGCGAGG ACCATCTTGA TGACACCGGC GACACCCGCG GCGGCCTGAG  
 27661 GTGTGGCTTC GTTGGACTTC AGCGAGCCCA GCAGACCGG GGTGTGCGCG CCCTGCCCGT  
 27721 AGGTGGCCAG CACCGCCTGT GCCTCGATGG GATCGCCAG CCTGGTGCCG GTGCGGTGCG  
 30 27781 CCTCCACGGC GTCCACGTCC GCGGGGTGA GCCCGGCGTT GGCCAGGGCC TGCCGGATCA  
 27841 CCGGCTCCTG CGAGGGCCCG TTCGGCGCCG ACAACCCGTT GGAAGCACCG TCCTGGTTGA  
 27901 CCGCCGAACC CCGGACAACC GCCAGCACAC GGTGGCCGTT GCGCTCGGCA TCGGAGAGCC  
 27961 TCTCGACGAT CAGCACACCG GACCCCTCGG CGAAACCGGT GCCGTCAGCC GCATCCJCGA  
 35 28021 ACGCCTTGCA GCGCGCGTCG GCGCGGAGC CCCGCTGCTG GGAGAACTCG ACGAAGCCCG  
 28081 ACGCGAGGC CATCACCGT ACGCCGCGA CGAGGGCGAG CGAGCATTCG CCGGAGCGCA  
 28141 GTGACTGCCC GGCCTGGTGC AGCGCCACCA GCGACGACGA ACACGCCGTG TCGACCGTGA  
 28201 CCGCCGGACC CTCCAGACCG TAGAAGTACG ACAGCCGACC GGACAGCACA CTGGTCTGGG  
 28261 TGCCGGTTCG GCCGAAACCG CCCAGGTCGG TGCCGAGTCC GTACCCGTCG GAGAAGGCGC  
 28321 CCATGAACAC GCCGGTGTG CTTCCGCGCA GCGACTCCGG GAGGATCCCG GCGTGTCCA  
 40 28381 GCGCCTCCCA CGAGGTCTCC AGGACCAGAC GCTGCTGCGG GTCCATCGCC AGCGCCTCAC  
 28441 GCGGACTGAT CCCGAAGAAC GCGCGTCGA AGTCCGCCAC CCCGGCGAGG AAGCCACCAT  
 28501 GACGACGGT CGACGTGCCC GGATGATCCG GATCGGGATC GTACGGGATC TCCAGTCCC  
 28561 AACCACGGTC CGTCGGAAC GCCGTGATCC CGTCACCACC CGACTCCAGC AGCCGCCACA  
 28621 AGTCTCCGG CGACGCGACC CACCCCGGCA GCCGGCAGGC CATCCCCACG ATCGCCAACG  
 45 28681 GTCGTCCTG CCGGACGGCC GCGGTCTGG TGCGGGTCGG CGATGCCGTC CGGCCGACA  
 28741 GCGCCGCGGT GAGCTTCGCC GCGACGGCGC GCGGCGTCGG GAAGTCGAAG ACCGCGGTGG  
 28801 CCGGCAGCCG TACGCCCGTC GCCTCGGTGA AGGCGTTGCG CAGCCGGATC GCCATGAGCG  
 28861 AGTCGACGCC GAGTTCCTTG AACGTGGCGG TCGCTCGAC CCGTGCGGCA CCGTCGTGGC  
 28921 CGAGTACGGC CGCGGTGCAC TGCCGGACGA CGGCGAGCAC GTCCTTTTCG GCGTCCGCGG  
 50 28981 CGGAGAGCCG CGCGATCCGG TCGGCGAGGG TGGTGGCGCC GGCCGCCCGG CGCCGCGGCT  
 29041 CCGGGCGCGG TGCGCGCAGC AGGGGCGAGC TGCCGAGGCC GGCCGGGTGCG GCGGCGACCA  
 29101 GCGCCGGGTC CGAGGACCGC AACGCCGCGT CGAACAGCGT CAGTCCGCTC TCGGCGJTC  
 29161 GCGCCGTAC GCCGTGCGCG CGCATGCGGG CGCCGGTGCC GACCGTCAGC CCGCTCTCCG  
 29221 GTTCCACAG GCCCAGGCC ACGGACAACG CGGGCAGTCC GGCTGCCCGG CGCTGTTCCG  
 55 29281 GAGGACGCGT GAGGAACGCG CTGCGGCGCG CGTAGTTGCC CTGCGGGG CTGCGGAGCA  
 29341 CACCGGCGGC CGACGAGTAG AGGACGAACG CGGCCAGTTC CGTGTCTGG GTGAGTTCGT  
 29401 GCAGGTGCCA CGCGGCGTCC ACCTTCGGGC GCAGACCGT CTCGAGCCGG TCGGGGGTGA  
 29461 GCGCGGTGAG GACGCCGTCG TCGAGGACGG CCGCGGTGTG CACGACGGCC GTGAGCGGGT  
 29521 GCGCCGGGTC GATCCCCGCC AGTACGGAGG CGAGTTCGTC CCGGTGCGCG ACCTCGCAGG  
 60 29581 CGATCGCCGT GACCTCGGCG CCGGGCACGT CGCTCGCCGT GCCGTGCGC GACAGCATCA  
 29641 GCAGCCGGCG CACGCCGTGG CGTTCGACGA GGTGGCGGCT GATGATGCCG GCCAGCGTCC  
 29701 CGGAGCCACC GGTGACGAGC ACGGTGCCGT CCGGGTCGAG CGCCGGAGCG TCACCCGCGG  
 29761 GGACCGCCGG GGCCAGACGG CGGGCGTACA CCTGGCCGTC ACGCAGCACC ACCTGGGGCT  
 29821 CATCGAGCGC GGTGGCCGCT GCGAGCAGCG GCTCGGCGGT GTCCGGGGCG GCGTCGACGA

29881 GGACGATCCG GCCGGGGTGT TCGGCCTGCG CGGTCCGCAC CAGTCCGGCG GCCGCGGCCG  
 29941 ACGCGAGACC GGGCCCCGTG TGGACGGCCA GGACCGCGTC GGCGTACCGG TCGTCGGTGA  
 30001 GGAAGCGCTG CACGGCGGTC AGGACGCCGG CGCCAGTTTC GCGGGTGTCTG TCGAGCGGGG  
 5 30061 CACCCCGCGC GCCGTGCGCG GGGAGGATCA CCACGTCCGG GACCGTCGGG TCGTCGAGGC  
 30121 GGCCGGTCTG CGCGGTCTGT GCGGCGAGCT CCGGGAGCTC GGCCAGCACC GGGCGCAGCA  
 30181 GGCCCGGAAC GGCTCCCGTG ATCGTCAGGG GCGCCTGCG CACGGCGCCG ATGGTGGCGA  
 30241 CGGGCCCCGC GTCTCTGTC GCGAGGTGTA CGCCGTCAGC GGTGACGGCG ACGCGTACCG  
 30301 CCGTGGCGCC GGTGGCGTGG GCGAGGTGTA CGCCGTCAGC GGTGACGGCG ACGCGTACCG  
 10 30361 CCGCGGCGAG GCGGAGTGCG GCGCCGAGCA GCGCCGGGTG CAGGCCGTAC CGTCCGGCGT  
 30421 CGGCGAGCTG TCCGTGCGCG AGGGCCACTT CCGCCAGAC GGCGTCGTCTG TCGGCCAGA  
 30481 CGGCGCGCGG GCGGGGCGAG GCGGGCCCGT CCGTGTACCC GGCTCGGGCC AGACGGTCGG  
 30541 CGATGTCTGT GGGGTCCACC GGCCGGGCGG TGGCGGGCGG CCACGTGAC GGCATCTCCC  
 30601 GCACGGCCGG GGCCGTCCCG GGTGCGGGG CGAGGATTCC GTGCGCGTGC TCGGTCCACT  
 15 30661 CCCCCCGCG GTGCCGCGTG TGACGGTGA CCGCGCGCG GCGGTCCGCC CCGGGCGGCC  
 30721 TCACCGTGAC GGAGAGCGCG AGCGCACCGG ACCGCGGAG CGTGAGGGG TGTCCACGG  
 30781 TGAACGTGTC GAGGGCGCGG CAGCGGCTT CGTCGCCCG CCGGATCGCC AGATCCAGGA  
 30841 GGCCCGCGGC GGGCAGCACC GCGAGGCCGT GCAGGGAGTG CGCCAGCGGA TCGGCGGCGT  
 30901 CGACCCGGCC GGTGAGCACC AGGTGCGCGG TGCCGGGCGG GGTGACCGCC GCGGTACGCG  
 20 30961 CCGGGTGCGC GACCGGCGTC TGTCGGGCGG GGGCCGCGTC GCGCGCGGTG TGGGTGCCGA  
 31021 GCGAGTAGCG GACCCGCTCG AACGGGTACG TCGGCGGGTG CGAGGCGCGT GCCGCGCGG  
 31081 GGTCTGATGAC CTTCGGCCAG TCGACCGTGA CGCCGTCTGG GTGCAGCCG GCGAGCGCGG  
 31141 TCAGGGCGGA TCGCGGTTCTG TCGTCGGCGT GCAGCATCGG GATGCCGTG ACGAGTCGGG  
 31201 TCAGGCTCCG GTCCGGGCGG ATCTCCAGGA GCACCGCCCC GTCGTGCGCG GCGACCTGTT  
 31261 CCCCAGACCG GACGGTGTCT GCGACCTGTC GTACCCAGTA CTCCGGCGTG GTGCAGGCGG  
 25 31321 CGCCCGCGGC CATCGGGATC CTCGGCTCGT GGTACGTGAG GCTCTCCCG ACCTTGCGGA  
 31381 ACTCCTCGAG CATCGGCTCC ATCCGCGCGG AGTGAACGC GTGGCTGGTC CGCAGGCGGG  
 31441 TGAAGCGGCC GAGCCGGGCG GCGACGTGTA GCACCGCCTC CTCGTACCG GAGAGCACGA  
 31501 TCGACGCGGG CCGGTTGACC CCGGCGATCT CCACGCCGTC CCGCAGCAGC CCGTCCGCGT  
 30 31561 CCGGTTCCGA CGCGATCAG GCGGCCATCG CCGCGCCGGA CGGCAGCGCC TGCATCAGGC  
 31621 GGGCCCGTGC GGACACCAGC CTGCACGCGT CCTCCAGGGA CCAGACGCCG GCGACGTACG  
 31681 CGGCGGCCAG CTCGCCGATC GAATGGCCCA CGAAGGCGTC CCGGCGTACG CCCCACGCCT  
 31741 CGAGCTGTGC GCCGAGTGCG ACCTGGAGCG CGAACACCGC GGGCTGGGCG TACCCGGTGT  
 31801 CGTGGAGGTC GAGCCCGGCG GGCAGTCTGA GGGCGTCCAG CACCTCGCG CGAGTGCGCG  
 35 31861 CGAAGACGTC GTAGGCGGCG GCGAGTCTGT GCGCCATGCC GGGACGTTGT GAGCCGTCG  
 31921 CGGAGAAGAG CCACACGAGG CCGCGGTCCG GTTCTGCGGC GCCGTTGACC GTGTCTGTC  
 31981 CGATCAGCGC GGCCCGGTGC GGGGAAGGCG TGCGGGCGAG CAGGGCCGCG GCCACGCGCG  
 32041 GCTCGTCCTC CTCGCCGGTG GCGAGGTGGG CGCGCAGGCG GTGTACCTGT GCGTCGAGTG  
 32101 CCTGCGGGGT GCGTGCCGAG AGCAGCAGGG GCAGCGGTCC GGTGTCTGGT GCCCGGGCGG  
 40 32161 GTTCGGGGGC CGGTGCGGGG TGGCTTTCTGA GGATGATGTG AGCGTTGGTG CCGCTAACCG  
 32221 CGAAGGAGGA CACCCCGGCG CGCGTGGGCG GGTGCGTTTC GGGCCAGGGG CCGGCGTCCG  
 32281 TGAGGAGTTC GACGGCGCGG CCGGTCTAGT GCAGGTGCGA GCGCGGCTG TCCAGTGCA  
 32341 GGGTTCGCGG CAGGGTGCCG TGCCGATGG CGAGGACCAT CTTGATGACA CCGCGACGC  
 32401 CCGCGGCGGC CTGAGTGTGG CCGATGTTGG ACTTCAGCGA GCGCAGCAGC ACCGGGGTGT  
 45 32461 CGCGATGCTG CCGTAGGTG GCCAGTACCG CCTGCGCCTC GATGGGGTCTG CCCAGCCTGG  
 32521 TCCCGGTGCC ATGCGCCTCG ACAGCGTCCA CATCCGCGCG GGTGAGCCCG GCGTTGGCCA  
 32581 GCGCCTGCCG GATCACCCTG TCCTGCGAGC GCGCGTTCTG CGCCGACAAC CCGTTGGAAG  
 32641 CACCGTCTG GTTGACCGCC GAACCACGCA CGACCGCCAG GACATTGTGG CCGTGCCGCT  
 32701 CGGCGTCCGA GAGCCTCTCG ACGATCAGCA CACCGGATCC CTCGCGGAAA CCGTGCCAT  
 50 32761 CAGCCGCATC CGCGAACGCC TTGACGCGG CGTCCGGGGA GAGGCCCGCG TGCTGGGAGA  
 32821 AGTCCACGAA GCCGGACGGC GAGGCCATCA CCGTGACGCC GCGGACCAGC GCGAGCGAGC  
 32881 ACTCCCCCGA GCGCAGCGAC TGCCCGGCCCT GGTGCGAGCG CACGAGCGAC GACGAACACG  
 32941 CCGTGTCCAC CGTGACCGCC GGACCCTCCA AACCGTAGAA GTACGACAGC CGACCGGACA  
 33001 GCACACTGGT CTGGGTGCTG GTGGCACCGA AACCGCCGCG GTCCGGTCCA GTGCCGACC  
 55 33061 CGTAGAAGTA GCCGCCCATG AACACGCCGG TGTCGCTTCC GCGCAGCGAC TCCGGGAGGA  
 33121 TCCCGGCGTG TTCCAGCGCC TCCACGAGG TCTCCAGGAC CAGACGCTGC TCGGAGTCCA  
 33181 TCGCCAGCGC CTCACGCGGA CTGATCCCGA AGAACGCGCG GTCGAAGTCC GCCACCCCGG  
 33241 CGAGGAAGCC ACCATGACGC ACGGTGACG TGCCCGGATG ATCCGGATCG GGATCGTACA  
 33301 GCGCGTCCAC GTCCCAACCA CCGTCCGTG GAAACGCCGT GATCCCGTCA CCACCGACT  
 60 33361 CCAGCAGCCG CCACAAGTCC TCCGGCGAG CGACCCACC CGGCAGCCGG CAGGCCATCC  
 33421 CCACGATCGC CAACGGCTCG TCCTGCCGGA CGGCGCGGT CCGGGTACG CGCGGGTGG  
 33481 TCGCCGCGCG GCCGGCCAGT TCGTCCAGT GCGCGCGAG CGCTGCGCT GCGCGTGGT  
 33541 CGAAGACGAG CGTAGCGGCG AGCGTCAGG CCGTCCGCTC GCGCGCGGT GCGATGGCGT  
 33601 CGACGCGCGT CAGCGAGTCG AAGCCACTT CCCTGAACGC GCGCGCGGT GCGATGGCGT  
 33661 GGGCGTCCG GTGGCCGAG ACCGCGCGAG CGCTGGTACG GACGAGGTG AGCATGTGCG

	33721	GCGCGGCCGG	AGGTGCGGAC	GTGCGCCGGA	CGGCCGGCAC	GAGGGTGCGT	AGGACCGGGC
	33781	GGACCCGGTC	GGACGCGGCG	ACGGCGGCGA	GGTCGAGCCG	GATCGGCACG	AGCGCGGGCC
	33841	GGTCGGTGTG	CAGGGCCGCG	TGCAACAGGG	CGAGCCCCTG	TGCGGCCGTC	ATCCGGGTCA
	33901	TGCCGTTGCG	GGCGATGCGG	GCCAGGTCCG	TGGCGGTCAG	CCGCCCGCCC	ATCCCGTCCG
5	33961	CCGCCTCCCA	CAGTCCCCAG	GCGAGCGAGA	CGGCGGGCAG	CCCCTGGTGG	TGCCGGTGGC
	34021	GGGCGAGCGC	GTCGAGGAAC	GCGTTGCCGG	TCGCGTAGTT	GGCCTGACCC	GCGCCGCCGA
	34081	ACGTGGCGGA	TATGGACGAG	TACAGGACGA	ACGCGGCCAG	GTCGAGATCG	CGCGTCAGCT
	34141	CGTGCAAGGTG	CCAGGCGACG	TCCGCCTTGA	CCCGCAGCAC	GGCGTCCCAC	TGCTCCGGCC
	34201	GCATGGTCTGT	CACGGCCGCG	TCGTGACGGA	TCCCGGCCAT	GTGCACGACG	GCGCGCAGCC
10	34261	GCTGGGCGAC	GTCGGGCGAG	ACTGCGGCCA	GCTCGTCGCG	GTCGACGACG	TCCGGCGCCA
	34321	CGTACCGCAC	GCGGTCTGTC	TCCGGCGTGT	CGCCGGGCGG	GCCGTTGCGG	GACACACAGA
	34381	CGACCTCGGC	GGCCTCGTGC	ACGGTGAGCA	GGTGGTCCAC	GAGGAGGCGG	CCGAGCCCGC
	34441	CGGTGCCGCC	GGTGACGAGG	ACGGTCCCGC	CGGTGACGCG	GGAGGTTCCG	GTGGCCGCGG
	34501	CGACACGGCG	CAGACGGGCC	GCACGCGCTG	TGCCGTCGGC	GACCCGGACG	TGCGGCTCGT
15	34561	CGCCGGCGGC	GAGCCCGGCC	GCTATGGCGG	CGGGCGTGAT	CTCGTCCGCT	TCGATCAGCG
	34621	CGACGCGGCC	GGGATGCTCC	GTCTCCGCGG	TCCGGACCAG	GCCGCCGAGC	GCTTCCGCGG
	34681	CGGGATCGCC	GGTACGGGTG	GCCACGATGA	GCCGGGATCG	CGCCAGCAGC	GGCTCGGCGA
	34741	GCCAGGTCTG	CACGGTGGTG	AGCAGGTGCG	GGCCGAGCTC	CCGGTCCGCG	GCGCCGGGCG
	34801	AGGTGCCCGG	GTCGCCGGGT	TCCACGGCCA	GGACCACGAC	CGGGGGGTGC	TCGCCGTCGG
20	34861	GCACGTGCGC	GAGGTACGTC	CAGTCGGGGA	CGGGTGACGC	GGGCACGGGC	ACCCAGGCGA
	34921	TCTCGAACAG	CGCCTCGGCA	TCGGGGTTCG	CGGCCCGCAC	GGTCAGGCTG	TCGACGTCAA
	34981	GGACCGGTGA	GCCGTGCTCG	TCCGTGGCGA	CGATGCGGAC	CATGTCGGGG	CCGACGCGTT
	35041	CCAGCAGCAC	GCGCAGCGCG	GTGCGGCGCG	GCGCGTGGAT	CCTCACGCGG	GACCAGGAGA
	35101	ACGCCAGCCG	GCGCCGCTCC	GGGTCCGTGA	AGACCGTCCC	GAGGCGTCCG	AGGGCCGCGT
25	35161	CGAGCAGCAC	GGGGTGACGC	CCGTACCGGG	CGTCGGTGAG	CTGTTCCGCG	AGGCCGACCG
	35221	ACGCGTAGGC	GCGGCCCTCC	CCCGTCCACA	TCGCGGTCAT	GGCCCGGAAC	GCGGGCCCGT
	35281	ACGAGAGCGG	CAGCGCGTCG	TAGAAGCCGG	TCAGGTGCGC	CGGGTCCGCG	TCGGCGGGCG
	35341	GCCAGTCCAC	GGGCTCCGCC	GGACCGCCAG	TGTCCACGCT	CAGCGTCCG	GTCCGACTGA
	35401	GCGCCCAAGG	GCCCCTGCCG	GTACGGCTGT	GCAGACTCAC	CGACCGCCGT	CCGGACACCT
30	35461	CGGTTCCGAC	GGTGGCCTGG	ATCTCCGTGT	CGCCGTGCGC	GTCGACACAC	ACCGCGCGGA
	35521	CGATGGTTCAG	CTCCGCGATC	TCCGCGTGC	CGAGCCGGGC	TCCCGCTTCG	TCGAGCAGTT
	35581	CCACGAGCGC	CGAGCCGGGC	ACGATGACCC	GGCCGTCCAC	CTCGTGGTGC	GCGAGCCAGG
	35641	GCTGACGGCG	TACCGAGACA	CCGCGGTGGC	CAGCGCGCCC	TCGCCGTCGG	GCGAGGTCTA
	35701	CCCACGAGCC	GAGCAGCGGG	TGGCCGGACG	TTCGCCCGCG	TTCGCCGTCG	ATCCAGTACC
35	35761	GGTCACGGCG	GAACGGGTAC	GTGGGCAGCG	GCACCACCCG	ACGCGTCCGC	AACGACGAGG
	35821	TGACGGGCAC	GCCCCGGACC	CAGAGCGCGG	CGAGCGACCG	AGTGAAGCGG	TCCAGGCCCG
	35881	CTCCGCTCGC	CCGCAGTGTG	CCGGTGACGA	CCGTATGCGC	ATGCCCGGCG	AGCGTGTCCCT
	35941	CCAGTGCGGT	GGTGAGCACG	GGATGCGCGC	TGACCTCGAC	GAACGCGCGG	TATCCGCGGT
	36001	CCGCCAGGTG	CCCGGTTCGG	GCGGCGAACC	GAACGGTGCG	GCGCAGGTTG	TCGTACCAGT
40	36061	AGGCGGCGTC	CGCGGGCCGG	TCCAGCCACG	CCTCGTCCAC	GGTGGAGAAG	AACGGGACGT
	36121	CCGGCGTGCG	CGGAGTGATG	CCGGCGAGAG	CGTCGAGCAG	CGCGCCGCGG	ATCGTTTCGA
	36181	CATGCGCGGT	GTGCGACGCG	TAGTCGACGG	CGATCCGGCG	GGCGCGGGGG	GTGGCGGCCA
	36241	GCAGTCTCTC	CACGGCGTCG	GCCGACCCGG	CGACAACGAT	CGACGCGGGT	CCGTTGACCG
	36301	CGCGCAGCTC	CAGGCGCCCC	CCCCACACGG	CGGCGTCGAA	GTCGCGGGGC	GGCAGCCGAG
45	36361	CCATGCCGCC	CTGCCCGGCC	AGTTCGGTGG	CGACGAGTCG	GCTGCGCACC	GCGACGACCT
	36421	TCGCGGCGTC	GTCCAGGGTG	AGCACCCCGG	CGACGCAGGC	CGCGGCGACT	TCGCCCTGGG
	36481	AGTGGCCGAC	GACCGCGGCC	GGGGCGACCC	CGTGCGCACG	CCACAGCTCC	GCCAGCGCCA
	36541	CCATCACCGC	GAACGACGCG	GGCTGCACGA	CATCGACCCG	GTCGAACGCG	GGCGCTCCGG
	36601	GCCGCTGGGC	GATGACGTCC	AGCAGGTCCC	ATCCGGTGTG	CGGGGCGAGC	GCCGTGGCGC
50	36661	ACTCGCGGAG	CCGCCGGGCG	AACACGGGCT	CGGTGGCGAG	CAGTTCGGCA	CCCATGCCCG
	36721	CCCACTGGGA	GCCCTGCCCG	GGGAACGCGA	ACACGACACG	TGTGTCGGTG	ACGTGCGCGG
	36781	TTCGCGTCAC	GGCCCCCGGC	ACTTCGGCAC	CACGGGCGAA	CGCCTCCGCC	TCTCGGGCCG
	36841	GCACGACCGC	CCGGTGGCGC	ATGGCCGTCC	GGGTGGTGGC	GAGCGAGTGG	CCGACCGCGG
	36901	CCGCGGCGCC	AGTGAGCGGG	GCCAGCTGTC	CCGCGACGTC	CCGCACTCCC	TCCGGGCTCC
55	36961	GGGCCGACAT	CGGCCAGACC	ACGTCTCTCG	GCACCGGCTC	GGCTTCGGGT	GCGGACACGG
	37021	GTGCGGGCGC	GGCGGGGGGC	CCGGCTTCCA	GGACGACATG	GGCGTTGGTG	CCGTGATGTC
	37081	GCAACGACGA	GACACCGGCA	CGCCGGTGGC	CGCCGCTGAC	CGGCCACGGC	TCACTGCGGT
	37141	GCAGCAGCCG	GATGTCGCGG	TCCAGTTCGA	CGTGCCGGGA	CGGCTCGTCG	ACGTGCAGCG
	37201	TGCGCGGCAG	GACGCCGTGC	CGCATCGCCA	TGACCATCTT	GATGACGCGG	GCGACGCCGG
60	37261	CCGCGGCCTG	GGTGTGGCCG	ATGTTCTGACT	TGAGCGAGCC	GATCAGCAGC	GGATGCACGC
	37321	GTTCGCGCCC	GTAGGCCACT	TGCAGGGCCT	GGGCCTCGAC	GGGGTCCGCC	AGACGGGTGC
	37381	CGGTGCCGTC	TGCCTCCACG	GCGTCGACGT	CACCCGGCGC	CAGGCCGGCG	TCGGCGAGCG
	37441	CACGCTGGAT	GACGCGCTGC	TGCGCAGGCC	CGTTCGGGGC	GGACAGCCCG	TTCGACGCGC
	37501	CGTCGGAGTT	GACCGCGGAG	CCGCGCACCA	GCGCCAGCAC	GGGGTGGCCG	TGGCGGGTGG



	37561	CGTCGGAGAG	CCGCTCCAGC	ACCAGGACAC	CGGCGCCCTC	GGCGAAGCTC	GTGCCGTCCG
	37621	CGGTGTCCGC	GAAGGCCTTG	GCACGGCCGT	CGGGGGCGAG	CCCGCGCTGC	CGGGAGAACT
	37681	CGACGAACCC	GGTCGTCTGC	GCCATCACCG	TGACACCGCC	GACCAGGGCG	AGCGAGCACT
	37741	CCCCCGAGCG	CAGCGACCGC	GCGGCCTGGT	GCAGCGCCAC	CAGCGACGAC	GAACACGCCG
5	37801	TGTCGACGGT	GACCGACGGG	CCCTCCAGAC	CGAAGTAGTA	CGAGAGCCGC	CCGGAGAGAA
	37861	CGCTGGTCGG	CGTGCCGGTC	GCCCCGAAAC	CGCCCAGGTC	CACGCCCGCG	CCGTAGCCCT
	37921	CGGTGAACGC	GCCCCATGAAT	ACGCCGGTGT	CGCTGCCGCG	GACGCTTTTCG	GGCAGGATGC
	37981	CCGCTCGTTC	GAACGCCCTC	CACGACGCTT	CGAGGACCAG	ACGCTGCTGC	GGGTCCATCG
10	38041	CCAGCGCCTC	ACGCGGGCTG	ATCCCAGAAG	ACGCGGCGTC	GAAGTCGCTC	GGCCCGGTGA
	38101	GGAAGCCGCC	GTGACGCACG	GAAACCTTGC	CGACCGCGTC	GGGGTTCGGG	TGCTAGAGCG
	38161	CGGCGAGGTC	CCAGCCGCGG	TCGGCGGGGA	ACTCGGTGAT	CGCGTCCCCG	CCGGAGTCGA
	38221	CCAGCCGCCA	CAGGTCTCTC	GGTGACCGCA	CGCCACCGGG	CATCCGGCAC	GCCATGGCCA
	38281	CGATCGCCAG	CGGCTCGTTC	CCCGCCACCG	TCGGTGCGGG	CACGTGCGCC	GCCGGAGCGG
	38341	CAGGGGCCGG	CTCACCCTCG	CGTTCCTCAT	CCAGGCGGGC	GGCGAGCGCG	GCCGGTGTCTG
15	38401	GGTGGTCGAA	GACGGCCGTC	CGGGAGAGCC	GTACCCCGT	CGTCTCGGCG	AGGTGTGTGC
	38461	GCAACCGGAC	ACCGCTGAGC	GAGTGCATGC	CGAGGTCTCT	GAACGCCGTC	GTGGCGGTGA
	38521	TCTCGGAGGC	GTGCGCGTGG	CCGAGCACGG	CGGCCGTGGC	CGCACACACG	ATGGCCAGCA
	38581	GGTCACGATC	GCGGTGCGGG	TCGCGGTGCG	GGTTGTCTCT	CGCACGGGCG	GCGATGCGGC
20	38641	GCTCGGTCCG	CTGCCGGACG	GGCTCGGTGG	GAATCGCCGC	GACCATGAAC	GGCACGTCCG
	38701	CGGCGAGGCT	CGCGTCGATG	AAGTGGGTGC	CCTCGGCCTC	GGTGAGCGGC	CGGAACCCGT
	38761	CGCGCACCCG	CTGCCGGTCG	GCGTCGTCAA	GTGTGTCGGT	GAGGGTGCTG	GTGGTGTGCC
	38821	ACATGCCCCA	GGCGATGGAG	GTGGCGGTTT	GGCCGAGGGT	GTGGCGGTGG	GTGGCGAGGG
	38881	CGTCGAGGAA	GGCGTTGGCG	GCGGCGTAGT	TTCTTGTCTC	GGGGCTGCCG	AGGACGGCGG
25	38941	CGGCGCTGGA	GTAGAGGACG	AAGTGGGTGA	GGGGTTGGTT	TTGGGTGAGG	TGGTGCAGGT
	39001	GCCAGGCGGC	GTTGGCTTTG	GGGTGGAGGA	CGGTGGTGAG	GCGGTGCGGG	GTGAGGGCGT
	39061	CGAGGATGCC	GTGCTCGAGG	GTGGCGGCGG	TGTGGAAGAC	GGCGGTGAGG	GGTTGGGGGA
	39121	TGTGGGCGAG	GGTGGTGGCG	AGTTGGTGGG	GGTCGCCGAC	GTCGCGAGGG	AGGTGGGTGC
	39181	CGGGGTGGT	GTGCGGGTGG	GGGGTGCGGG	AGAGGAGGTA	GGTGTGCGGG	TGGTTCAGGT
30	39241	GGCGGGCGAG	GATGCCGGCG	AGGGTGCCGG	AGCCGCCGGT	GATGATGATG	GCGTGTTCGG
	39301	GGTTGAGGGG	GGTGGTGGTG	GGTGGGGTGG	TGGTGTGGAG	GGGGGTGAGG	TGGGTTCGGT
	39361	GGAGGGTGTG	GTGGGTGAGG	CGGAGGTGGG	GGTGGTCGAG	GGTGGCGAGT	TGGGCCAGGG
	39421	GGAGGGGAGT	GTGGGGGTGG	TCGGTTTTCG	TGAGGCGGAT	GCGGTGGGGG	TGTTCTGTTCT
	39481	GGGCGGTGCG	GGTGAGGCCG	GTGACGGTGG	CGCCGCGCGG	GTCGGTGGTG	GTGTGGACGA
35	39541	TGAGGGTGTG	GTGCGTGGTG	GTGAGGTGGT	GTGTCAGGGC	GGTCAGGACG	CGGGTGGCGC
	39601	GGGTGTGGGC	GCGGGTGGGT	ATGTCCTCGG	GGTCGTGCGG	GTGGGCGGCG	GTCATCAGGA
	39661	CGTGTCCCTC	GGGCAGGTCA	CCGTCTGAGA	ACCGCTCGGC	GACCGGAGC	GATCCCAACC
	39721	GGAGCGGGTT	CGGCCCGGAC	GGGGTGTGCG	CCCGCTCCCT	CAGCACCAGC	GAGTCCACCG
	39781	ACACGACAGG	ACGGCCATCC	GGGTGCGCCA	CGCGCACGGC	GACGCCGGCC	TCCCCCGGGG
40	39841	TGAGGGCGAC	GCGCACCGCG	GCGGCCCGCG	TGGCGTTCAG	GCGCACGCCC	GTCCAGGAGA
	39901	ACGGCAGCTC	GATCCCGCGC	CCCGCGTCTG	GGCGCCCGGC	GTGCAGGGCC	GCGTCGAGCA
	39961	GTGCCGGATG	CACACCGAAA	CCGTCCGCCCT	CGGCGGCCCT	CTCGTCGGGC	AGCGCCACCT
	40021	CGGCATACAC	GGTGTACCA	TCACGCGAGG	CAGCCCGCAA	CCCCTGGAAC	GCCGACCCGT
	40081	ACTCATAACC	GGCATCCCGC	AGTTCGTCTAT	AGAACCCCGA	GACGTGACAG	GCCGCGGCCG
45	40141	TGGCCGGCGG	CCACTGCGAG	AACGGCTCAC	CGGAAGCGTT	GGAGGTATCC	GGGGTGTCCG
	40201	CGGTACAGGT	GCCGCTGGCG	TGCCGGGTCC	AGCTGCCCCT	GCCCTCGGTA	CGCGCTGGA
	40261	CGGTACCCGG	CCGCCGTCCG	GCCTCATCGG	CCCTTCCAC	GGTCACCGAC	ACATCCACCG
	40321	CTGCGGTAC	CGGCACCACG	AGCGGGGATT	CGATGACCAG	TTCATCCACC	ACCCCGCAAC
	40381	CGGTCTCGTC	ACCGGCCCGG	ATGACCAGCT	CCACAAACGC	CGTACCCGGC	AGCAGAACCG
50	40441	TGCCCCGCAC	CGCGTGATCA	GCCAGCCAGG	GATGCGTACG	CAATGAGATC	CGGCCGGTGA
	40501	GAACAACACC	ACCACCGTCG	TCGGCGGGCA	GTGCTGTGAC	GGCGGCCAGC	ATCGGATGCG
	40561	CCGCCCCGGT	CAGCCCGGCC	GCGGACAGGT	CGGTGGCACC	GGCCGCCCTCC	AGCCAGTACC
	40621	GCTGTGCTC	GAACGCGTAG	GTGGGCAGAT	CCAGCAGCCG	CCCCGGCACC	GGTTCGACCA
	40681	CCGTGCCCCA	GTCCACCCCC	GCACCCAGAG	TCCACGCCCT	CGCCAACGCC	CCCAGCCACC
55	40741	GCTCCAGCC	ACCGTCACCA	GTCCGCAACG	ACGCCACCGT	GCGGGCCTGT	TCCATCGCCC
	40801	GCAGCAGCAC	CGATGGGCA	CTGCACTCCA	CGAACACCGA	CCCCTCCAGC	TCCGCCACCG
	40861	CCGCATCCAG	CGCGACAGGG	CGACGACAGT	TCCGGTACCA	GTACCCGCTC	TCCACCGGCT
	40921	CGGTACCCCA	GGCGCTGTCC	ACGGTCGACC	ACCACGCCAC	CGACCCGGTC	CCGCCGGAAA
	40981	TTCCCTTCAG	TACCTCAGCG	AGTTCGTCTT	CGATGGCCTC	CACGTGAGGC	GTGTGGGAGG
60	41041	CGTAGTCGAC	CGCGATACGA	CGCACCCGCA	CCCCATCAGC	CTCATACCGC	GCCACCACCT
	41101	CCTCCACCGC	CGACGGGTCC	CCCGCCACCA	CCGTGGAAGC	CGGACCATTA	CGCGCCGCGA
	41161	TCCACACACC	CTCGACCAGA	CCCACCTCAC	CGGCCGGCAA	CGCCACCGAA	GCCATCGCCC
	41221	CCCGCCCGGC	CAGCCGCGCC	GCGATCACCC	GACTGCGCAA	CGCCACCACG	CGGGCCGCGT
	41281	CCTCCAGGCT	GAGGGCTCCG	GACACACACG	CCGCCGCGAT	CTCCCCCTGC	GAGTGTCCGA
	41341	CCACAGCGTC	CGGCACGACC	CCATGCGCCT	GCCACAGCGC	GGCCAGGCTC	ACCGCGACCG



41401 CCCAGCTGGC CGGCTGGACC ACCTCCACCC GCTCCGCCAC ATCCGACCGC GACAACATCT  
 41461 CCCGCACATC CCAGCCCGTG TGCGGCAACA ACGCCCGCGC ACACTCCTCC ATACGAGCCG  
 41521 CGAACACCGC GGAACGGTCC ATGAGTTCCA CGCCCATGCC CACCCACTGG GCACCCTGCC  
 41581 CGGGGAAGAC GAACACCGTA CGCGGCTGAT CCACCGCCAC ACCCATCACC CGGGCATCAC  
 5 41641 CCAGCAGCAC CGCACGGTGA CCGAAGACAG CACGCTCACG CACCAACCCC TGCGCGACCG  
 41701 CGGCCACATC CACCCACCCC CGCGCAGAT ACCCCTCCAG CCGCTCCACC TGCCCCCGCA  
 41761 GACTCACCTC ACCACGAGCC GACACCGGCA ACGGCACCAA CCCATCACCA CCGACTCCA  
 41821 CACGCGACGG CCCAGGAACA CCTCCAGGA TCACGTGCGC GTTCGTACCG CTCACCCCGA  
 41881 ACGACGACAC ACCCGCATGC GGTGCCCGAT CCGACTCGGG CCACGGCCTC GCCTCGGTGA  
 10 41941 CGAGTCCAC CGCACCGGCC GACCACTCCA CATGCGACGA CGGCTCGTCC ACGTGCAGCG  
 42001 TCTTCGGCGC GATCCCATGC CGCATCGCCA TGACCATCTT GATGACACCG CGGACACCCG  
 42061 CAGCCGCCTG CGCATGACCG ATGTTCTGACT TGACCGAACC GAGGTAGAGC GCGGTGTGCG  
 42121 GGTCTTGCCC GTAGGCCGCG AGGACGGCCT GCGCCTCGAT CGGGTCGCCC AGCCGCGTGC  
 42181 CGGTGCCGTG CGCCTCCACC ACGTCCACAT CGGCGGCGCG CAGTCCGGCG TTGACCAACG  
 15 42241 CCTGCCGGAT CACGCGCTGC TGGGCGACGC CGTTGGGGGC GGACAGTCCG TTGGAGGCAC  
 42301 CGTCTGGTT CACCGCCGAG CCGCGAGACA CGCGGAGAAC GGTGTGCCCC TTGCGCTCGG  
 42361 CGTCGGAGAG CCGCTCCAGC ACGAGAACGC CGACGCCCTC GGCGAAGCCG GTCCCGTCCG  
 42421 CCGCGTCGGC GAACGCCTTG CACCGTCCGT CCGGGGAGAG TCCGCGCTGC CGGGAGAACT  
 42481 CCACGAGCTC TGCGGTGTTT GCCATGACGG TGACACCGCC GACCAGCGCC AGGGAGCACT  
 20 42541 CCCCAGCCCG CAGTGCCTGT GCGCCTGGT GCAGGGCGAC CAGCGACGAC GAGCAGCCG  
 42601 TGTCGACCGT GACCGCCGGG CCTGAAGTC CGTACACGTA CGAGAGGCGC CCGGACAGGA  
 42661 CGTCTGCTG CGTCGCCGTG ACACCGAGCC CGCCAGGTC CCGGCCGACG CCGTAGCCCT  
 42721 GGTGAACGC GCCATGAAC ACGCCGGTGT CGCTCTCCCG GAGCCTGTCC GGCACGATCG  
 42781 CCGCGTTCTC GAACGCCTCC CAGGAGGTCT CCAGGATCAG GCGCTGTGG GGTCCATCG  
 25 42841 CCAGCGCCTC GTTCGACTG ATGCCAAGA ACGCGGCGTC GAACCCGGCG CCGGCCAGGA  
 42901 ATCCGCCGTG GCGTGTGCTG GAGCGGCCG CCGCGTCCGG GTCCGGGTG TACAGCGCGT  
 42961 CGACGTCCCA GCCCCGGTCG GTGGGGAAC CCGGTGATCGC CTCGGTACCG GCGGCGACGA  
 43021 GCGCCACAG GTCTCCCGG GAGGCGACCC CGCCGGGCG TCGGCACGCC ATGCCGACGA  
 43081 TCGCGACGGG GTCGCCGGAG CCGAGGGTCT GGGCGGTGCG GGTGCGCGT GTCCGGGAGC  
 30 43141 CCGCGAGGTG GCGCGGAAC GCACGCGGAG TGGGGTGGTC GAACGCGGTT GCGCGGGCA  
 43201 CCGCAGACC CGTCCGCGCG GCGACGGTGT TGGTGAACTC GACGGTGGTG AGCGAGTCGA  
 43261 GGCCGTTCTC GCGGAACGTG CCGTCCGGGG AGCAGTGTCC GGCGCCCGGC AGGCCAGGA  
 43321 CCGTGCGGAC GCTGTGCGCG ACCAGGTCGA GCAGTACGTC CTCCCGGCC GCACGGGCCG  
 43381 CCGCGAGGCG GTTCGCCAC TCCTGTTCCG TGGCGTCGGG CTCGGCCGGT CCGGTCAGTG  
 35 43441 CCGTGAGGAT CCGCGGCGTG GCGCCCGCCA TCGTCGCGG CCGCGCCCGG TCGGAACCGG  
 43501 TCCGGGCCAC GATGTACGAG CGCGCCCGCG CGATGGCCTT CTCGATCAGG GCGCGGCTCA  
 43561 GCGCCGGCCG TTCGATGCCG GGCAGCGCGC GGACGGTGAC GGTGGGGAGT CCCTCCGCGG  
 43621 CCGGTGGCCG GGTGTGGCG TCGGCGCCGG CCGGGCCGTC GAGCAGGACG TGCACGAGCG  
 43681 CGCCGGGGTT CGCGGCTTCC TCGGCTGCGG TGGTACAGTG GGTGAGGCCG GTCTCGTGC  
 40 43741 GGAGCAGGCC GCGGACGGTG TCGGCGTCTT CCCCCTGAC CAGGACCGGC GCGTCCGGGC  
 43801 CGATCGGAGG CGGCACGGTG AGGACCATCT TGCCGGTGTG CCGGGCGTGG CTCATCCACG  
 43861 CGAACGCGTC CCGCGCACGG CCGATGTCCC ACGGTGTCAC CCGCAGCGGG CACAGCTCAC  
 43921 CGCGGTCGAA CAGGTGAGG AGCAGTTCGA GGATCTCCCG CAGGCGCGG GGTCCACGT  
 43981 CCGCCAGGTC GAACGGCTGC TGGGCGGCGT GGCGGATGTC GGTCTTGCCC ATCTCGACGA  
 45 44041 ACCGCGCGCC CCGTGCGAGC AGGCCGATGG ACGCGTCGAG GAGTTCACCG GTGAGCGAGT  
 44101 TGAGCACGAC GTCGACCGGC GGAAGGTGT CCGCGAACGC GCGCTGCGG GAGTTCGCCA  
 44161 CATGGTCGGT GTCGAAGCCG TCGGCGTGCA GCAGGTGTTG TTTGGCGGGA CTGGCGGTGG  
 44221 CGTACACCTC GCGCGCGAGG TGGCGGGCGA TCCGGGTGCG CGCCATGCCG ACACCGCCCG  
 44281 TCGCGGCGTG GACCAAGACC TTCTGGCCCG GTCGCGAGTC GCCCGGTCG ACGAGGCCGT  
 50 44341 ACCAGCGGT GCGGAACACG ATGGGACCGG ACGCGCGGAT GGGGAACGAC CATCCCCGTG  
 44401 GGATCCGTGC GACCAGCCGC CCGTCCGCGA CCACGCTGCG CCGGAACGCG TCCTGCACGA  
 44461 GACCGAACAC GCGGTGCGCG GGGGCCAGGT CGTCGACGCC GGGTCCGACT TCGGTACGA  
 44521 TGCCCGCGGC CTCCCCGCC ATCTCGCCCT CGCCCGGGTA GGTGCCGAGC GCGATCAGCA  
 44581 CGTCGCGGAA GTTCAGCCCC GCGGCGCGGA CGTCGATGCG GACCTCGCCG GCGGCCAGGG  
 55 44641 GCGCGGCGGG ACGTCGAGCG GGGCGACGAG GAGGTGCGCG AGCGTTCGCG AGGCGGGCGG  
 44701 CGCGAGCGCC CACTGGCGCG GTGCGGAGGG GGGTGGTGTG CCGCGCTACC AGCCGGGCA  
 44761 CGTAGGCCAC GCCGGCCCCG AGCGCGATCT GGGGTTGCGC GAGCGAGGCC GCGGCGGGGA  
 44821 CGAGGTGCTC ATCGCCGTCC GTGTCCACCA GCACGAACGA TCCGGGTTG GCGGCTGGC  
 44881 GGCGCAGCGC CTCGTCCCAG AGCCGGGCCCT GGTCCGCGTC CCGGATCTCG GCCGGGCCGA  
 60 44941 CGCCACCGC GCGGCGGGTG ACGACCGTCC GCGGGGTTGA CCGGGTGCCG GGCAGGTGCG  
 45001 GCGGCTCCCA GACCAAGTTC CACAGCGTGG CCTCGCCACT GCGGTTGGCG ACCAGATGGG  
 45061 CCGGAGCGCC CGCGAGCCG CGCGCTGGA CTTTGCCCGA CCGGTTGCGG GGGATCGTGG  
 45121 TGACGTGCCA GATCTGTCG GGCACCTTGA AGTAGGCGAG CCGGCGCGG CACTCGGCGA  
 45181 GGATCGCCTC GCGGGGACG CCGGGGCCGT CGGAAACGAC GTAGAGCAGG GGTATGTGCG

45241 CGAGGACGGG GTGCGGGCGG CCCGCCGCGG CGGCGTCCCG GACACCGGCC ACCTCCTGGG  
45301 CGACGGTCTC GATCTCCCGG GGGTGGATGT TCTCCCGGCC GCGGATGATC AGCTCCTTGA  
45361 CCCGGCCGGT GATCGTCACG TGTCCGGTCT CGGCCTGACG TGCGAGGTCC CCGGTGCGGT  
5 45421 ACCAGCCGTC CACGAGCACC TGGGCGGTCT CCTCCGGCTG GGCGTGGTAG CCGAGCATGA  
45481 GGCTCGGCC GCTCGCCAC AGCTCGCCCT CCTCGCCGGG TGCCACGTCT GCGCCGGACA  
45541 CCGGGTCGAC GAACCGCAGC GACAGGCCCG GCACGGGCAG CCCGCACGAG CCGGGAACCC  
45601 GCGCATCCTC CAGGGTGTG GCGGTGAGCG AGCCGGTCTG CTCGGTGCAG CCGTACGTGT  
45661 CGAGCAGGGG CACGCCGAAC GTCGCTCGA AATCCCTGGT GAGCGACGCC GGCGAGGTGG  
10 45721 ATCCGGCGAC CAGCGCCACG CGCAGCGCGC GAGCCCGCGG CTCGCCGGAC ACGGCGCCGA  
45781 GGAGGTAGCG GTACATCGTC GGCACGCCGA CGAGCACGGT GCTGGAGTGT TCGGCCAGGG  
45841 CGTCGAGGAC GTCACGCGCG ACGAAGCCGC CCAGGATACG GGCGGACGCG CCGACCGTGA  
45901 GGACGGCGAG CAGGCAGAGG TGGTGGCCGA GGCTGTGGAA CAGCGGGGCG GGCCAGAGCA  
45961 GTTCGTCTG CTCGGTCAG CGCCAGGACG GCACGTGCA GTGCATCGCG GACCAAGCG  
15 46021 CGCTGCGCTG TCGGAAACC ACGCCCTTGG GACGGCCGGT GGTGCCGGAG GTGTAGAGCA  
46081 TCCAGGCGGG TTCGTCCAGG CCGAGGTCTG CGCGGGGCGG GCACGGCGGC TCGGTCCCGG  
46141 CGAGGTCTCT GTAGGAGACG CAGTCCGGTG CCCGGCGCCC GACGAGCACG ACGGTGGCGT  
46201 CCGTGCCGGT GCGGCGCACC TGGTCGAGGT GGGTTTCGTC GGTGACCAGC ACGGTGCGCG  
46261 CGGAGTCCCT CAGGAAGTGG GCGAGTTCGG CGTCGGCGGC GTCCGGGTTG AGCGGGACCG  
20 46321 CGACGCGCGG GCGCGGGCG GCGCGGAGT AGACCTCGAT GGTCTCGATC CGGTTGCCGA  
46381 GCAGCATGCG GACCCGGTCG CCGCGGTCGA GCGCGGACGC GCGGAGGTGT CCGGCGAGCC  
46441 GGCCGGCCCG GAGCCGGAGT TCGGTGTACG TCACGGCGCG TTGGGAATCC GTGTAGGCGA  
46501 TCCGGTCCCG GCGTCGCTCG GCATGGATGC GGAGCAATTC GTGCAACGGC CGGATTGGTT  
46561 CCACACGCGC CATGGAAACA CCTTCTCTC GACCAACCGC ACAACAGCAC GGAACCGGCC  
25 46621 ACGAGTAGAC GCCGGCGACG CTAGCAGCGT TTTCCGGACC GCCACCCCTT GAAGATCCCC  
46681 CTACCGTGGC CGGCCTCCCC GGACGCTCAT CTAGGGGGTT GCACGCATAC CGCCGTGCGT  
46741 AATTGCCTTC CTGATGACCG ATGCCGGACG CCAGGGAAGG GTGGAGGCGT TGTCCATATC  
46801 TGTCACGGCG CCGTATTGCC GCTTCGAGAA GACCGGATCA CCGGACCTCG AGGGTGACGA  
46861 GACGGTGCTC GGCCTGATCG AGCACGGCAC CGGCCACACC GACGTGTGCG TGGTGGACGG  
30 46921 TGCTCCCCGG ACCGCCGTGC ACACCACGAC CCGTGACGAC GAGGCGTTCA CCGAGGTCTG  
46981 GCACGCACAG CGCCCTGTCT AGTCCGGCAT GGACAACGGC ATCGCCTGGG CCCGCACCGA  
47041 CGCGTACCTG TTCGGTGTCT TGCGCACCGG CGAGAGCGGC AGGTACGCCG ATGCCACCGC  
47101 GGCCCTCTAC ACGAACGTCT TCCAGTCTAC CCGGTCGCTG GGGTATCCCC TGCTCGCCCG  
47161 GACCTGGAAC TACGTCAGCG GTATCAACAC GACGAACGCG GACGAGTCTG AGGTGTACCG  
35 47221 GGACTTCTGC GTGGGCCGCG CCCAGGCGCT CGACGAGGGC GGGATCGACC GCGCCACCAT  
47281 GCGCGCGGCC ACCGGTATCG GCGCCACGG GGGCGGCATC ACCTGCGTGT TCCTCGCCGC  
47341 CCGGGGCGGA GTCGGATCA ACATCGAGAA CCCC GCCGTC CTCACGGCCC ACCACTACCC  
47401 GACGACGTAC GGTCCGCGGC CCCC GGCTCTT CGCACGGGCC ACCTGGCTGG GCGCGCCGCA  
47461 GGGGGGCGCG CTGTTCTCT CCGGACGGC CGGCATCCTC GGACACCGAA CCGTGCACCA  
40 47521 CCGTGATGTC ACCGGCCAGT GCGAGGTGCG CCTCGACAAC ATGGCCCGCG TCATCGGCGC  
47581 GGAGAACCTG CCGCGCCACG GCGTCCAGCG GGGGCGACGC CTCGCCGACG TGACCACTT  
47641 CAAGGTCTAC GTCCGCCGCC GCGAGGATCT CGATACGGTC CGCCGGGTCT GCGCCGACG  
47701 CCTGTGAGC ACCGCGGCCG TCGCCCTTTT GCACACCGAC ATAGCCCGCG AGGATCTGCT  
47761 CGTCGAAATC GAAGGCATGG TGGCGTGACA ATACCCGGTA AAAGGCCCGC GACGCTGCGC  
45 47821 CTCGGCGGAT CCGCGAAGAG AAAGAAGAGC GTCACCGCAC AGCGCGGCAG CCCGGTCTT  
47881 TCGTCTCTCG CACAGCGGCG GATCTGGTTT CTCCAGCAAT TGGACCCGGA GAGCAACGCC  
47941 TATAATCTCC CGTCTGTGCA ACGCTGCGC GGTCTATTGG ACGCGCCGGC CCTGGAGCGT  
48001 GCGCTGGCGC TCGTCTGTCG GCGCCACGAG CGGTGCGGA CCGTGTTCGA CACCGCCGAC  
48061 GGCGAGCCCC TCCAGCGGGT GCTTCCCGCC CCGGAACACC TCCTGCGCCA GCGCGGGCG  
50 48121 GGCAGCGAGG AGGACGCCGC CCGGCTCGTC CGCGACGAGA TCGCCGCGCC GTTCGACCTC  
48181 GCCACCGGGC CGTTGATCAG GGCCCTGCTG ATCCGCCTCG GTGACGACGA CCACGTTCTC  
48241 GCGGTGACCG TGCACCATGT CGCCGGCGAC GGCTGGTCTG TCGGGCTCCT CCAACATGAA  
48301 CTCGCAGCCC ACTACAGGCG GCTGCGCGAC ACTGCCCGCC CTGCCGAAT GCGCCGTTG  
48361 CCGGTGACG ACGCCGACTT GCGCGCTGG GAGCGGCGCG AACTCACC GGCGGACTG  
55 48421 GACAGGCGTC TGGCCTACTG GCGCGAGCAA CTCCGGGGCG CCGCGGCGCG GCTCGCCCTC  
48481 CCCACCGACC CTCCCCGCC GCGGTGCGC GACGCGGACG CCGGCATGGC CGAGTGGCGG  
48541 CCGCGGCGCG CGCTGGCCAC CGCGGTCTCT ACGCTCGCGC GCGACTCCGG TCGTCCGTG  
48601 TTCATGACCC TGCTGGCGGC CTTCCAAGCG GTCCTCGCCC GGCAGGCGGG CACGCGGAC  
48661 GTGCTGGTCT GCACGCCCGT GGCGAACCGT ACGCGGGCGG CGTACGAGGG CCTGATCGGC  
60 48721 ATCTTCGTCA ACACGCTCGC GCTGCGCGCG GACCTCTCGG GCGATCCGTG GTTCCGGGAA  
48781 CTCTTCGACC GCTGCCGGCG CACGACACG GACGCGTTG CCCACCGCA CCGCGCTTC  
48841 GAGAACGTCA TCGAACTCGT CGCACCGGAA CGCGACCTGT CCGTCAACCC GGTGCTCCAG  
48901 GTGCTGTTGC AGGTGCTGCG GCGCGACGCG GCGACGGCG CGCTGCCCGG CATCGCGGCC  
48961 GAACCGTTCC GCACCGGACG CTGGTTCACC CGCTTCGACC TCGAATTCCA TGTGTACGAG  
49021 GAGCCGGGTG GCGCGCTGAC CGCGAACTG CTCTACAGCC GTGCGCTGTT CGACGAGCCA

49081 CGGATCACGG GGTGCTGGA GGAGTTCACG GCGGTGCTTC AGGCGGTCAC CGCCGACCCG  
 49141 GACGTACGGC TGTCGCGGCT GCCGGCCGGC GACGCGACGG CGGCAGCGCC CGTGGTGCCC  
 49201 TCGAACGACA CGGCGCGGGA CTTGCCGTC GACACGCTGC CGGGCCTGCT GGCCCGGTAC  
 49261 GCCGCACGCA CCCCCGGCGC CGTGCGCGTC ACCGACCCGC ACATCTCCCT CACCTACGCG  
 5 49321 CAGCTGGACC GGCGGGCGAA CCGCTCGCG CACCTGCTCC GCGCGCGCGG CACCGCCACC  
 49381 GGCGACCTGG TCGGGATCTG CGCGGATCGC GGCGCCGACC TGATCGTGGG CATCGTGGGG  
 49441 ATCCTCAAGG CGGGCGCCGC TTATGTGCCG CTGGACCCCG AACATCCTCC GGAGCGCACG  
 49501 GCGTTCGTGC TGGCCGACGC GCAGCTGACC ACGGTGGTGG CGCACGAGGT CTACCGTTCC  
 49561 CGGTTCCTCC ATGTGCCGCA CGTGGTGGCG TTGGACGACC CGGAGCTGGA CCGGACCCG  
 10 49621 GACGACACGG CGCCGGACGT CGAGCTGGAC CGGGACAGCC TCGCCTACGC GATCTACACG  
 49681 TCCGGGTCGA CCGGCAGGCC GAAGGCCGTG CTCATGCCGG GTGTACGCGC CGTCAACCTG  
 49741 CTGCTCTGGC AGGAGCGCAC GATGGGCCGC GAGCCGGCCA GCCGACCGT CCAGTTCGTG  
 49801 ACGCCACGT TCGACTACTC GGTGCAGGAG ATCTTTTCCG CGCTGCTGGG CGGCACGCTC  
 49861 GTCATCCCGC CGGACGAGGT GCGGTTTCGAC CCGCCGGGAC TCGCCCGGTG GATGGACGAA  
 15 49921 CAGGCGATTA CCCGATCTA CGCGCCGACG GCCGTACTGC GCGCGCTGAT CGAGCACGTC  
 49981 GATCCGACA GCGACAGCT CGCCGCCCTG CGGCACCTGT GCCAGGCGG CGAGGCGCTG  
 50041 ATCCTCGACG CGCGGTTGCG CGAGCTGTGC CGGCACCGGC CCCACCTGCG CGTGACAAT  
 50101 CACTACGGTC CGGCCGAAAG CCAGCTCATC ACCGGGTACA CGCTGCCCGC CGACCCCGAC  
 50161 GCGTGCCCG CCACCGCACC GATCGGCCCG CCGATCGACA ACACCCGCAT CCATCTGCTC  
 20 50221 GACGAGGCGA TGCGGCCGGT TCCGGACGGT ATGCCGGGGC AGCTCTGCGT CGCCGGCGTC  
 50281 GGCTCGCCC GTGGGTACCT GGCCCGTCCC GAGCTGACCG CCGAGCGCTG GGTGCCGGGA  
 50341 GATCGGTGCG GCGAGGAGCG CATGTACCTC ACCGGCGACC TGGCCCGCCG CGCGCCCGAC  
 50401 GGCGACCTGG AATTCTCGG CCGGATCGAC GACAGGTCA AGATCCGCGG CATCCGCGTC  
 50461 GAACCGGGTG AGATCGAGAG CCTGCTCGCC GAGGACGCCC GCGTCACGCA GCGCGCGGTG  
 25 50521 TCCGTGCGCG AGGACCGGCG GGGCGAGAAG TTCTGGCCG CGTACGTCGT ACCGGTGGCC  
 50581 GGCCGGCACG GCGACGACTT CGCCGCGTCG CTGCGCGCGG GACTGGCCGC CCGGCTGCCC  
 50641 GCCGCGCTCG TGCCCTCCGC CGTCGTCCTG GTGGAGCGAC TGCCGAGGAC CACGAGCGGC  
 50701 AAGGTGGACC GGCGCGCGCT GCCCGACCCG GAGCCGGGCC CGGCGTCGAC CGGGGCGGTT  
 50761 ACGCCCCGCA CCGATGCCGA CGGACGGTG TGCCGATCT TCCAGGAGT GCTCGACGTC  
 30 50821 CCGCGGGTCG GTGCCGACGA CGACTTCTT ACGCTCGGCG GGCATCCCT GCTGCCACC  
 50881 CGGGTCGTCT CCCGATCCG CGCCGAGCTG GGTGCCGATG TCCGCTGCG TACGCTCTTC  
 50941 GACGGGCGGA CGCCCGCCGC GCTCGCCGT GCGGCGGACG AGGCCGGCCC GGCCGCCCTG  
 51001 CCCCCGATCG CGCCCTCCGC GGAGAACGGG CCGGCCCCCC TCACCGCGGC ACAGGAACAG  
 51061 ATGCTGCACT CGCACGGCTC GCTGCTCGCC GCGCCCTCCT ACACGGTCGC CCCGTACGGG  
 35 51121 TTCCGGCTGC GCGGGCCACT CGACCGCGAA CGGCTCGACG CGGCACTGAC CCGGATCGCC  
 51181 CGCGGCCACG AGCCGCTGCG GACCGGTTT CCGCATCGGG AACAGTCTGT CCGCCCGCCC  
 51241 GCTCCGGTGC GCGCCGAGGT GGTTCGGTG CCGGTCGGCG ACGTCGACGC CGCGGTCCGG  
 51301 GTCGCCACC GGGAGCTGAC CCGGCCGTT GACCTCGTGA ACGGTCGTT GCTGCGTGCC  
 51361 GTGCTGCTGC CGCTGGGCGC CGAGGATCAC GTGCTGCTGC TGATGCTGCA CCACCTCGCC  
 40 51421 GGTGACGGAT GGTCTTTCGA CCTCTGGTC CCGGAGTTGT CCGGGACGCA ACCGGACCTT  
 51481 CCGGTGTCTT ACACGGACGT GGCCCGGTGG GAACGGAGTC CGGCCGTGAT CGCGGCCAGG  
 51541 GAGAACGACC GGGCCTACTG GCGCCGCGCG CTGGGGGGCG CCACCGGCC CGAGCTGCC  
 51601 GCGGTCCGGC CCGGCGGGGC ACCGACGGG CGGGCGTTCC TGTGGACGTC CAAGGACACC  
 51661 GCGTCTCTGG CGGCACGCCG GGTGCGGGAC GCCCACGACG CGACGTTGCA CGAAACCGTG  
 45 51721 CTCGGCGCCT TCGCCCTGGT CGTGGCGGAG ACCGCCGACA CCGACGACGT GCTCGTCGCG  
 51781 ACGCCGTTTC CGGACCGGGG GTACGCGGG ACCGACCACC TCATCGGCTT CTTGCGGAAG  
 51841 GTCCTCGCGC TGCGCCTCGA CCTCGGCGGC ACGCCGTCTG TCCCGAGGT GCTGCGCCGG  
 51901 CTGCACACCG CGATGGTGGG CGCGCACGCC CACGAGGCGG TGCCCTACTC CGCGCTGCGC  
 51961 GCCGAGGACC CCGCGCTGCC GCCGCCCCC GTGTCGTTCC AGCTCATAG CGCGCTCAGC  
 50 52021 CGGGAACGTC GGCTGCCCCG CATGCACACC GAGCCGTTCC CCGTCGTCGC CGAGACCGTC  
 52081 GACGAGATGA CCGGCGAACT GTGATCAAC CTCTTCGACG ACGGTCGCAC CGTCTCCGGC  
 52141 GCGGTGGTCC ACGATGCCGC GCTGCTCGAC CGTGCCACCG TCGACGATTT GCTCACC CGG  
 52201 GTGGAGGCGA CGCTGCGTGC CGCCGCGGGC GACCTCACC TACGCGTCAC CGGTTACGTG  
 52261 GAAAGCGAGT AGCCATGCCC GAGCAGGACA AGACAGTCGA GTACCTTCGC TGGGCGACCG  
 55 52321 CGGAACCTCA GAAGACCGCT CGGGAACCTC CCGCGCACAG CGAGCCGTTG GCGATCGTGG  
 52381 GGTGCGCTG CCGGCTGCCC GCGGGGTCG CGTCGCCGGA GGACCTGTGG GAGTTGCTGG  
 52441 AGTCCGGTGG CGACGGCATC ACCGCGTTCC CCACGAGCCG GGGCTGGGAG ACCACCGCCG  
 52501 ACGGTCGCGG CGGCTTCCTC ACCGGGGCGG CCGGCTTCGA CGCGGCGTTC TTCGGCATCA  
 52561 GCCCGCGGGA GCGCTGGCG ATGGACCGGC AGCAGCGCCT GGCCCTGGAG ACCTCGTGGG  
 60 52621 AGGCGTTCGA GCACGCGGGC ATCGATCCGC AGACGCTGCG GGGCAGTGAC ACGGGGGTGT  
 52681 TCCTCGGCGC GTTCTTCCAG GGTACGGCA TCGGCGCCGA CTTGACGGT TACGGACCA  
 52741 CGAGCATTC CACGAGCGTG CTCTCCGGCC GCCTCGGTA CTTCTACGGT CTGGAGGCTC  
 52801 CGCGGTCAC GGTGACACG CGTGCTGGT GCGCTGTCAC GGCCTGTCAC GAGCCGGGC  
 52861 AGTCGCTGCG CTCCGGCGAA TGCTCGCTCG CCCTGGTCGG CGGCGTCACG GTGATGGCCT

	52921	CGCCGGCGGG	GTTCGCGGAC	TTCTCCGAGC	AGGGCGGCCT	GGCCCCGAC	GCGCGCTGCA
	52981	AGGCCTTCGC	GGAAGCGGCT	GACGGCACC	GTTTCGCCGA	GGGGTCCGGC	GTCTGATCG
	53041	TCGAGAAGCT	CTCCGACGCC	GAGCGCAACG	GCCACCGCGT	GCTGGCGGTC	GTCCGGGGTT
5	53101	CCGCCGTCAA	CCAGGACGGT	GCCTCCAACG	GGCTGTCCGC	GCCGAACGGG	CCGTGCGAGG
	53161	AGCGGGTGAT	CCGGCAGGCC	CTGGCCAACG	CCGGACTCAC	CCCGGCGGAC	GTGGACGCCG
	53221	TCGAGGCCCA	CGGCACCGGC	ACCAGGCTGG	GCGACCCCAT	CGAGGCACAG	GCCGTGCTGG
	53281	CCACCTACGG	GCAGGGGCGC	GACACCCCTG	TGCTGCTGGG	CTCGCTGAAG	TCCAACATCG
	53341	GCCACACCCA	GGCCGCGCGC	GGCGTCGCCG	GTGTCATCAA	GATGGTCCTC	GCCATGCGGC
10	53401	ACGGCACCCT	GCCCCGCACC	CTGCACGTGG	ACACGCCGTC	CTCGCACGTC	GA CTGGACGG
	53461	CCGGCGCCGT	CGAACTCCTC	ACCGACGCCC	GGCCCTGGCC	CGAAACCGAC	CGCCACGGC
	53521	GCGCCGGTGT	CTCCTCCTTC	GGCGTCAGCG	GCACCAACGC	CCACATCATC	CTCGAAAGCC
	53581	ACCCCGGACC	GGCCCCGAA	CCCGCCCCG	CACCCGACAC	CGGACCGCTG	CCGCTGCTGC
	53641	TCTCGGCCCG	CACCCGCGAG	GCATCGAGT	CACAGGTACA	CCGCTGCGC	GCGTTCCTCG
15	53701	ACGACAACCC	CGGCGCGGAC	CGGGTCGCCG	TCGCGCAGAC	ACTCGCCCGG	CGTACCCAGT
	53761	TCGAGCACCG	CGCCGTGCTG	CTCGGCGACA	CGCTCATCAC	CGTGAGCCCG	AACGCCGGCC
	53821	GCGGACCGGT	GGTCTTCGTC	TACTCGGGGC	AAAGCACGCT	GCACCCGCAC	ACCGGGCGGC
	53881	AACTCGCGTC	CACCTACCCC	GTGTTCCGCC	AAGCGTGGCG	CGAGGCCCTC	GACCACCTCG
	53941	ACCCACCCA	GGGCCCGGCC	ACGCACTTCG	CCCACCAGAC	CGCGCTCACC	GCGTCTCTGC
20	54001	GGTCTGGGG	CATACCCCG	CACGCGGTCA	TCGGCCACTC	CCTCGGTGAG	ATCACCGCCC
	54061	CGGACGCCG	CGGTGTCCTG	TCCCTGAGGG	ACGCGGGCGC	GCTCCTCACC	ACCCGCACCC
	54121	GCCTGATGGA	CCAAGTGCCG	TCGGGCGGCG	CGATGGTCAC	CGTCTTGACC	AGCGAGGAAA
	54181	AGGCAAGCCA	GGTGCTGCGG	CCGGGCGTGG	AGATCGCCGC	CGTCAACGGC	CCCCACTCCC
	54241	TCGTGCTGTC	CGGGGACGAG	GAAGCCGTAC	TCGAAGCCGC	CCGGCAGCTC	GGCATCCACC
25	54301	ACCGCCTGCC	GACCCGCCAC	GCCGGCCACT	CCGAGCGCAT	GCAGCCACTC	GTGCCCCCCC
	54361	TCCTCGACGT	CGCCCGGACC	CTGACGTACC	ACCAGCCCCA	CACCGCCATC	CCCGGCGACC
	54421	CCACCACCGC	CGAATACTGG	GCGCACCAGG	TCCGCGACCA	AGTACGTTTC	CAGGCGCACA
	54481	CCGAGCAGTA	CCCGGGCGCG	ACGTTCTCTG	AGATCGGCC	CAACCAGGAC	CTCTCGCCGC
	54541	TCGTGACGG	CGTTGCCGCC	CAGACCGGTA	CGCCCGACGA	GGTGCGGGCG	CTGCACACCG
30	54601	CGCTCGCGCA	GCTCCACGTC	CGCGGCGTCG	CGATCGACTG	GACGCTCGTC	CTCGGCGGGG
	54661	ACCGCGCGCC	CGTCACGCTG	CCCACGTATC	CGTTCCAGCA	CAAGGACTAC	TGGTTCGGGC
	54721	CCACCTCCCG	GGCCGATGTG	ACCGGCGCGG	GGCAGGAGCA	GGTGGCGCAC	CCGCTGCTCG
	54781	GCGCCGCGGT	CGCGCTGCCC	GGCACGGGCG	GAGTCGTCTT	GACCGGCCGC	CTGTGCTGCG
	54841	CCTCCCATCC	GTGGCTCGGC	GAGCAGCGCG	TCGACGGCAC	CGTGCTCCTG	CCCGGCGCGG
35	54901	CCTTCTCGA	ACTCGCGGCG	CGCGCCGGCG	ACGAGGTCGG	CTGCGACCTG	CTGCGACGAA
	54961	TCGTGATCGA	GACGCGGCTC	GTGCTGCCCC	CGACCGGCGG	TGTGGCGGTC	TCCGTGCGAG
	55021	TCGCCGAACC	CGACGACACG	GGGCGGCGGG	CGGTACCCGT	CCACGCGCGG	GCCGACGGCT
	55081	CGGGCCTGTG	GACCCGACAC	GCCGGCGGAT	TCCTCGGCAC	GGCACC GGCA	CCGGCCACGG
	55141	CCACGGACCC	GGCACCCTGG	CCGCCCGCGG	AAGCCGGACC	GGTCGACGTC	GCCGACGTCT
40	55201	ACGACCGGTT	CGAGGACATC	GGGTACTCCT	ACGGACCGGG	CTTCCGGGGG	CTGCGGGGCG
	55261	CCTGGCGCGC	CGGCGACACC	GTGTACGCCG	AGGTGCGGCT	CCCCGACGAG	CAGAGCGCCG
	55321	ACGCCGCCCG	TTTCACGCTG	CACCCCGCGC	TGCTCGACGC	CGCGTCCAG	GCCGGCGCGC
	55381	TGGCCGCGCT	CGACGCACCC	GGCGGGGCGG	CCCGACTGCC	GTTCTCGTTC	CAGGACGTCC
	55441	GCATCCACGC	GGCCGGGGCG	ACGCGGCTGC	GGGTACGGT	CGGCCGCGAC	GGCGAGCGCA
45	55501	GCACCGTCCG	CATGACCGGC	CCGGACGGGC	AGCTGGTGGC	CGTGGTGGGT	GCCGTGCTGT
	55561	CGCGCCCGTA	CGCGGAAGGC	TCCGGTGACG	GCCTGCTGCG	CCCGGTCTGG	ACCGAGCTGC
	55621	CGATGCCCGT	CCCGTCCGCG	GACGATCCCG	GCGTGGAGGT	CCTCGGCGCC	GACCCGGGCG
	55681	ACGGCGACGT	TCCGGCGGCC	ACCCGGGAGC	TGACCGCCCG	CGTCTCTCGG	GCGCTCCAGC
	55741	GCCACCTGTC	CGCCGCGGAG	GACACCACCT	TGGTGGTACG	GACCGGCACC	GGCCCGGCCC
50	55801	CTGCCGCGCG	CGCGGGTCTG	GTCCGCTCGG	CGCAGGCGGA	GAACCCCGGC	CGCGTCGTGC
	55861	TCGTGAGGCG	GTCCCGGAC	ACCTCGGTGG	AGCTGCTCGC	CGCGTGCGCC	GCGCTGGACG
	55921	AACCGCAGCT	GGCCGTCCGG	GACGGCGTGC	TCTTCGCGCC	GCGGCTGGTC	CGGATGTCCG
	55981	ACCCCGCGCA	CGGCCCGCTG	TCCCTGCCCG	ACGGCGACTG	GCTGCTCACC	CGGTCCGCCT
	56041	CCGGCACGTT	GCACGACGTG	GCGCTCATAG	CCGACGACAC	GCCCCGGCGG	GCGCTCGAAG
55	56101	CCGGCGAGGT	CCGATCGAC	GTCCGCGCGG	CCGGACTGAA	CTTCCGCGAT	GTGCTGATCG
	56161	CGCTCGGGAC	GTACACCGGG	GCCACGGCCA	TGGGCGGCGA	GGCCGCGGGC	GTGCTGGTGG
	56221	AGACCGGGCC	CGGCGTGGAC	GACCTGTCCC	CCGGCGACCG	GGTGTTCGGC	CTGACCCGGG
	56281	GCGGCATCGG	CCCGACGGCC	GTCACCGACC	GGCGCTGGCT	GGCCCGGATC	CCCGACGGCT
	56341	GGAGCTTCAC	CACGGCGGCG	TCCGTCCCGA	TCGTGTTCGC	GACCGCGTGG	TACGGCCTGG
60	56401	TCGACCTCGG	CACACTGCGC	GCCGGCGAGA	AGGTCTCTCG	CCACGCGGCT	ACCGCGGGTG
	56461	TGAGCATGGC	CGCCGCACAG	ATCGCCCGCC	ACCTGGGCGC	CGAGCTGACC	CGCACC GCA
	56521	GTACCGGCAA	GCAGCACGTC	GTCGCGCGCG	CCGGGCTGCC	CGACACGCAC	ATCGCCGACT
	56581	CTCGGACGAC	CGCGTTCCGG	ACCGCTTTCC	CGCGCATGGA	CGTCTCTCTG	AACGCGCTGA
	56641	CCGGCGAGTT	CATCGACGCG	TCGCTCGACC	TGCTGGACGC	CGACGGCCCG	TTGCTGCGAG
	56701	TGGGCGGCAC	CGAGCTGCGC	GACCCGGCCG	CGATCGTCCC	GCGCTACCTG	CCGTTCGACC

56761 TGCTGGACGC GGGCGCCGAC CGCATCGGCG AGATCCTGGG CGAACTGCTC CGGCTGTTCC  
 56821 ACGCGGGCGC GCTGGAGCCG CTGCCGGTCC GTGCCTGGGA CGTCCGGCAG GCACGCGACG  
 56881 CGCTCGGCTG GATGAGCCG GCCCGCCACA TCGGCAAGAA CGTCCTGACG CTGCCCCGGC  
 56941 CGCTCGACCC GGAGGGCGCC GTCGTCTCA CCGGCGGCTC CGCGCATCC  
 5 57001 TCGCCCGCCA CCTGCGCGAA CGGCATGTCT ACCTGCTGTC CCGGACGGCA CCGCCCGAGG  
 57061 GGACGCCCCG CGTCCACCTG CCCTGCGACG TCGGTGACCG GGACCAGCTG GCGGCGGCC  
 57121 TGGAGCGGGT GGACCGGCCG ATCACC GCCG TGGTGCACCT CGCCGGTGCG CTGGACGACG  
 57181 GCACCGTCGC GTCGCTCACC CCGAGCGTT TCGACACGGT GCTGCGCCCG AAGGCCGACG  
 57241 GCGCCTGGTA CCTGCACGAG CTGACGAAGG AGCAGGACCT CGCCGCGTTC GTGCTCTACT  
 10 57301 CGCTCGGCGC CGGCGTGCTC GGCAACGCCG GCCAGGGCAA CTACGTCCG CCGAACGCGT  
 57361 TCCTCGACGC GTCGCGCGAG CTGCGCCACG GTTCCGGGCT GCCGGCCCTC TCCATCGCCT  
 57421 GGGGGCTCTG GGAGGACGTG AGCGGGCTCA CCGGCGGCTC CGGCGAAGCC GACCGGGACC  
 57481 GGATGCGGCG CAGCGGTTTC CCGGCCATCA CCGCGCAACA GGGCATGCAC CTGTACGAGG  
 57541 CCGCCGGCCG CACCGGAAGT CCGGTGGTGG TCGCGGCGGC GCTCGACGAC GCGCCGGACG  
 15 57601 TGCCGCTGCT GCGCGGCCTG CCGCGGACGA CCGTCCGGCG GGCCGCCGTC CCGGAGTGTT  
 57661 CGTCCGCCGA CCGGCTCGCC GCGCTGACCG GCGACGAGCT CGCCGAAGCG CTGCTGACGC  
 57721 TCGTCCGGGA GAGCACGCC CCGGTGCTCG GCCACGTGGG TGGCGAGGAC ATCCCCGCGA  
 57781 CCGCGGCGTT CAAGGACCTC GGCATCGACT CGCTACCGC GGTCCAGCTG CGCAACGCC  
 57841 TCACCGAGGC GACCGGTGTG CCGCTGAACG CCACGGCGGT CTTGACTTC CCGACCCGCG  
 20 57901 ACGTGCTCGC CCGGAAGCTC GCGGACGAAC TGACCGGCAC CCGCGCGCCC GTCGTGCCCC  
 57961 GGACCGCGGC CACGGCCGGT GCGCACGACG AGCCGCTGGC GATCGTGGGA ATGGCCGTC  
 58021 GGCTGCCCGG CCGGGTCCGC TCACCCGAGG AGCTGTGGCA CCTCGTGGCA TCCGGCACCG  
 58081 AGCCCATCAC GGAGTTCCCG ACCCGACCG GCTGGGACGT CGACCGCATC TACGACCCG  
 58141 ACCCCGACGC GATCGGCAAG ACCTTCGTCC GGCACGGTGG CTTCTCACC GCGCGACAG  
 25 58201 GCTTCGACGC GCGGTTCTTC GGCATCAGCC CGCGCGAGGC CCTCGCGATG GACCCGACG  
 58261 AGCGGGTGCT CCTGGAGACG TCGTGGGAGG CGTTCGAAAG CGCCGGCATC ACCCCGGACT  
 58321 CGACCCGCGG CAGCGACACC GCGGTGTTCC TCGGCGCCTT CTCCTACGGT TACGGCACCG  
 58381 GTGCGGACAC CGACGGCTTC GCGCGACCG GCTCGCAGAC CAGTGTGCTC TCGGCGCGG  
 58441 TGTCGTACTT CTACGGTCTG GAGGGTCCGG CCGTCACGGT CGACACGGCG TGTTCTGCTG  
 30 58501 CGCTGGTGGC GCTGCACGAG CCGGGGACGT CGCTGCGCTC CGGCGAATGC TCGCTGCC  
 58561 TGGTCGGCGG CGTCACGGTG ATGGCGTCTC CCGGCGGCTT CGTGGAGTTC TCCCGCGAGC  
 58621 GCGGCCTCGC GCGGACGGC CCGGCGAAGG CGTTCGGCGG GGGTGGCGAC GGCACGAGCT  
 58681 TCGCCGAGGG TGCCGGTGTG CTGATCGTCG AGAGGCTCTC CGACGCCSAA CGCAACGGTC  
 58741 ACACCGTCCT GGCGGTGCTC CGTGGTTCCG CCGTCAACCA GGATGGTGCC TCCAACGGGC  
 35 58801 TGTCGGCGCC GAACGGGCGG TCGCAGGAGG GGGTGATCCG GCAGGCCCTG GCCAACCGCG  
 58861 GGCTACCCC GCGGACGTC GACGCGTCCG AGGCCACCG CACCGGCACC AGGCTGGGCG  
 58921 ACCCCATCGA GGCACAGCG GTACTGGCCA CCTACGGACA GGAGCGCGCC ACCCCCTGCG  
 58981 TGCTGGGCTC GCTGAAGTCC AACATCGGCC ACGCCAGGC CCGCTCCGGC GTCGCCGGCA  
 59041 TCATCAAGAT GGTGCAGGCC CTCGGGACG GGGAGCTGCC GCCGACGCTG CACGCCGACG  
 40 59101 AGCCGTCGCC GCACGTCGAC TGGACGGCCG GCGCCGTCGA ACTGCTGACG TCGGCCGCGC  
 59161 CGTGGCCCGA GACCGACCGG CCACGGCGTG CCGCCGTCTC CTCGTTCCGG GTGAGCGGCA  
 59221 CCAACGCCCA CGTCATCCTG GAGGCGGGAC CCGTAACGGA GACGCCCGCG GCATCGCCTT  
 59281 CCGGTGACCT TCCCCTGCTG GTCTCGGCAC GCTACCGGA AGCCTCGAC GAGCAGATCC  
 59341 GCCGACTGCG CGCCTACCTG GATCACCACC CCGACGTCGA CCGGGTGCCG GTGGCACAGA  
 45 59401 CGCTGGCCCG GCGCACACAC TTCGCCACC GCGCCGTGCT GCTCGGTGAC ACCGTCATCA  
 59461 CCACACCCCC CCGGACCGG CCGACGAAC TCGTCTTCGT CTACTCCGGC CAGGGCACCC  
 59521 AGCATCCCGC GATGGGCGAG CAGCTCGCCG CCGCCCATCC CGTGTTCGCC GACGCCTGGC  
 59581 ATGAAGCGCT CCGCCGCCTT GACAACCCCG ACCCCACGA CCCCACGCAC AGCCAGCATG  
 59641 TGCTCTTCGC CCACAGGCG GCGTTACCG CCCTCCTGCG GTCCTGGGGC ATACCCCGCG  
 50 59701 ACGCGGTCAT CCGCCACTCG CTGGGCGAGA TCACCGCGGC GCACGCGCGC GGCATCCTGT  
 59761 CGCTGGACGA CGCGTGACCC CTGATACCA CGCGCGCCCG CCTCATGCAC ACGCTCCCGC  
 59821 CACCCGGTGC CATGGTCACC GTACTGACCA GCGAAGAGAA GGCACGCCAG GCGTTGCGGC  
 59881 CCGGCGTGGA GATCGCCGCC GTCAACGGGC CCCACTCCAT CGTGCTGTCC GGGGACGAGG  
 59941 ACGCCGTGCT CACCGTCGCC GGGCAGCTCG GCATCCACCA CCGCCTGCCC GCCCCGACG  
 55 60001 CCGGGCACTC CGCGCACATG GAGCCCGTGG CCGCCGAGCT GCTCGCCACC ACCCGCGGGC  
 60061 TCCGCTACCA CCCTCCCCAC ACCTCCATT CCGAACGCC CACCAACGCT GAGTACTGGG  
 60121 CCGAGCAGGT CCGCAAGCCC GTGCTGTTCC ACGCCACGC CAGACGATAC CCGGACGCCG  
 60181 TGTTCTGTGA GATCGGCCCC GCGGAGACC TCTCCCGCT CGTCGACGGG ATCCCGCTGC  
 60241 AGAACGGCAC CGCGGACGAG GTGCACGCGC TGCACACCGC GCTCGCGCAC CTCTACGCGC  
 60 60301 GCGGTGCCAC GCTCGACTGG CCGCGCATCC TCGGGGCTGG GTCACGGCAC GACGCGGATG  
 60361 TGCCCGCGTA CGCGTTCCAA CCGCGGCACT ACTGGATCGA GTCGGCACGC CCGGCCGCAT  
 60421 CCGACGCGGG CACCCCGGTG CTGGGCTCCG GTATCGCCCT CGCCGGGTG CCGGGCGGGG  
 60481 TGTTACGGG TTCCGTGCCG ACCGGTGCGG ACCGCGCGGT GTTCGTGCGC GAGTGGCGC  
 60541 TGGCCGCGC GGACGCGGTC GACTGCGCCA CCGTCGAGCG GCTCGACATC GCCTCCGTGC

	60601	CCGGCCGGCC	GGGCCATGGC	CGGACGACCG	TACAGACCTG	GGTCGACGAG	CCGGCCGGACG
	60661	ACGGCCGGCG	CCGGTTACCC	GTGCACACCC	GCACCGGCGA	CGCCCCGTGG	ACGCTGCACG
	60721	CCGAGGGGGT	GCTGCGCCCC	CATGGCACGG	CCCTGCCCGA	TGCGGCCGAC	GCCGAGGGC
	60781	CCCCACCGGG	CGCGGTGCCC	GCGGACGGGC	TGCCGGGTGT	GTGGCGCCGG	GGGGACCAGG
5	60841	TCTTCGCCGA	GGCCGAGGTG	GACGGACCGG	ACGGTTTCGT	GGTGCACCCC	GACCTGCTCG
	60901	ACGCGGTCTT	CTCCGCGGTC	GGCGACGGAA	GCCGCCAGCC	GGCCGGATGG	CGCGACCTGA
	60961	CGGTGCACGC	GTCCGACGCC	ACCGTACTGC	GCGCCTGCCT	CACCCGGCGC	ACCGACGGAG
	61021	CCATGGGATT	CGCCGCCTTC	GACGGCGCCG	GCCCTGCCGT	ACTCACCCCG	GAGGCGGTGA
	61081	CGCTGCGGGA	GGTGGCGTCA	CCGTCCGGCT	CCGAGGAGTC	GGACGGCCTG	GACCGGTTGG
10	61141	AGTGGCTCGC	GGTCGCCGAG	GCGGTCTACG	ACGGTGACCT	GCCCCGAGGA	CATGTCCTGA
	61201	TCACCGCCGC	CCACCCCGAC	GACCCCGAGG	ACATACCCAC	CCGCGCCAC	ACCCGCGCCA
	61261	CCCGCGTCCT	GACCGCCCTG	CAACACCACC	TCACCACCAC	CGACCACACC	CTCATCGTCC
	61321	ACACCACCAC	CGACCCCGCC	GGCGCCACCG	TCACCGGCCT	CACCCGCACC	GCCCAGAACG
	61381	AACACCCCCA	CCGCATCCGC	CTCATCGAAA	CCGACCACCC	CCACACCCCC	CTCCCCCTGG
15	61441	CCCAACTCGC	CACCCCTCGAC	CACCCCTCAC	TCCGCCTCAC	CCACCACACC	CTCCACCACC
	61501	CCCACCTCAC	CCCCCTCCAC	ACCACCACCC	CACCCACCAC	CACCCCCCTC	AACCCCGAAC
	61561	ACGCCATCAT	CATCACCGGC	GGCTCCGGCA	CCCTCGCCGG	CATCCTCGCC	CGCCACCTGA
	61621	ACCACCCCCA	CACCTACCTC	CTCTCCCGCA	CCCCACCCCC	CGACGCCACC	CCCGGCACCC
	61681	ACCTCCCTTG	CGACGTGCGC	GACCCCCACC	AACTCGCCAC	CACCTCACCC	CACATCCCCC
20	61741	AACCCCTCAC	CGCCATCTTC	CACACCGCCG	CCACCCCTCGA	CGACGGCATC	CTCCACGCCC
	61801	TCACCCCGGA	CGCCTCACC	ACCGTCTCTC	ACCCCAAAGC	CAACGCCCGC	TGGCACCTGC
	61861	ACCACCTCAC	CCAAAACCAA	CCCCCTACCC	ACTTCGTCTC	CTACTCGAGC	GCCGCCJCCG
	61921	TCCTCGGCAG	CCCCGGACAA	GGAAACTACG	CCGCCGCCAA	CGCCTTCCTC	GACGCCCTCG
	61981	CCACCCACCG	CCACACCCCTC	GGCCAACCCG	CCACCTCCAT	CGCCTGGGGC	ATGTGGCACA
25	62041	CCACCAGCAC	CCTCACCGGA	CAACTCGACG	ACGCCGACCG	GGACCGCATC	CGCCCGGGCG
	62101	GTTTCCTCCC	GATCACGGAC	GACGAGGGCA	TGCGCCTCTA	CGAGGCGGCC	GTCCGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCCGGCACA	GCCGATGACC	GGCTCCGTAC
	62221	CGCCCATCCT	GAGCGGCCTG	CGCAGGAGCG	CGCGGCGCGT	CGCCCGTGCC	GGGACACCGT
	62281	TCGCCCAGCG	GCTCGCCGAG	CTGCCCGACG	CCGACCGCGG	CGCGGCGCTG	ACCCAGCTCG
30	62341	TCTCGGACGC	CACGGCCGCC	GTGCTCGGCC	ACGCCGACGC	CTCCGAGATC	GCGCCGACCA
	62401	CGACGTTCAA	GGACCTCGGC	ATCGACTCGC	TCACCGCGAT	CGAGCTGCGC	AACCGGCTCG
	62461	CGGAGGCGAC	CGGGCTGCGG	CTGAGTGCCA	CGCTGGTGTT	CGACCACCCG	ACACCTCGGG
	62521	TCCTCGCCGC	CAAGCTCCGC	ACCGATCTGT	TCGGCACGGC	CGTGCCACAG	CCCGCGCGGA
	62581	CGGCACGGAC	CCACCACGAC	GAGCCACTCG	CGATCGTCGG	CATGGCGTGC	CGACTGCCCC
35	62641	GCGGGGTCCG	CTCGCCGGAG	GACCTGTGGC	AGCTCGTGCG	GTCGCCACCC	GACGCGATCA
	62701	CCGAGTTCCC	CACCGACCGC	GGCTGGGACA	TCGACCGGCT	GTTTCGACCCG	GACCCGGACG
	62761	CCCCCGGCAA	GACCTACGTC	CGGCACGGCG	GCTTCCTCGC	CGAGGCCGCC	GGCTTCGATG
	62821	CCGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT	GGACCCGCAG	CAGCGCGTCA
	62881	TCCTCGAAAC	CTCCTGGGAG	GCGTTCGAGA	ACGCGGGCAT	CGTGCCGGAC	ACGCTGCGCG
40	62941	GCAGCGACAC	CGGCGTGTTT	ATGGGCGCGT	TCTCCCATGG	GTACGGCGCC	GGCGTCGACC
	63001	TGGGCGGGTT	CGGCGCCACC	GCCACGCAGA	ACAGCGTGCT	CTCCGGCCGG	TGTCGTTACT
	63061	TCTTCGGCAT	GGAGGGCCCG	GCCGTCACCG	TGCACACCGC	CTGCTCGTCG	TCGCTGGTCG
	63121	CCCTGCACCA	GGCGGCACAG	GCGCTGCGGA	CTGGAGAATG	CTCGCTGGCG	CTCGCCGGCG
	63181	GTGTACGGT	GATGCCCCACC	CCGCTGGGCT	ACGTCGAGTT	CTGCCGCCAG	CGGGGACTCG
45	63241	CCCCCGACGG	CCGTTGCCAG	GCTTCGCGGG	AAGGCGCCGA	CGGCACGAGC	TTCTCGGAGG
	63301	GCGCCGGCGT	TCTTGTTGCTG	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TCGCGGTCTG	CCGCTCCTCC	GCCGTCAACC	AGGACGGCGC	CTCCAACGGC	ATCTCCGCAC
	63421	CCAACGGCCC	CTCCCAGCAG	CGCGTCATCC	GCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
	63481	CCGCCGACGT	GGACGTGGTG	GAGGCCACAG	GCACCGGAAC	CCCGCTGGGC	GACCCGATCG
50	63541	AGGCACAGGC	CATCATCGCG	ACCTACGGCC	AGGACCGCGA	CACACCGCTC	TACCTCGGTT
	63601	CGGTCAAGTC	GAACATCGGA	CACACCCAGA	CCACCGCCGG	TGTCGCCCGC	GTCATCAAGA
	63661	TGGTCATGGC	GATGCGCCAC	GGCATCGCGC	CGAAGACACT	GCACGTGGAC	GAGCCGTCTG
	63721	CGCATGTGGA	CTGGACCGAG	GGTGCGGTGG	AACTGCTCAC	CGAGGCGAGG	CCGTGGCCCC
	63781	ACGCGGGACG	CCCGCGCCCG	GCGGCGGTGT	CGTCGCTCGG	TATCAGCGGT	ACGAACGCC
55	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	GCGGTGTGGA	GCCGCTGTTT	GCCGCTTGG
	63901	TGCCGTTGCC	GGTGTGCGCT	CGGAGTGAGG	CGAGTCTGCG	GGGGCAGGTG	GAGCGGCTGG
	63961	AGGGGTATCT	GCGCGGGAGT	GTGGATGTGG	CCGCGGTGCG	GCAGGGGTTG	GTGCGTGAGC
	64021	GTGCTGTCTT	CGGTACCCGT	GCGGTACTGC	TGGGTGATGC	CCGGGTGATG	GGTGTGGCGG
	64081	TGGATCAGCC	CCGTACGGTG	TTGCTCTTTC	CCGGGCAGGG	TGCTCAGTGG	GTGGGCATGG
60	64141	GTGTGGAGTT	GATGGACCGT	TCTGCGGTGT	TCGCGGCTCG	TATGGAGGAG	TGTGCGCGGG
	64201	CGTTGTTGCC	GCACACGGGC	TGGGATGTGC	GGGAGATGTT	GGCGCGGCCG	GATGTGGCGG
	64261	AGCGGGTGA	GGTGGTCCAG	CCGGCCAGCT	GGGCGGTGCG	GCTCAGCCTG	GCCGACTGT
	64321	GGCAGGCCCA	CGGGGTGCTA	CCCAGCGCGG	TGATCGGACA	CTCCCAGGGC	GAGATCGCGG
	64381	CGGCGTGCGT	GGCCGGGGCC	CTCAGCCTTG	AGGACGCCCG	CCGCGTGGTG	GCCTTGCGCA

	64441	GCCAGGTCAT	CGCGGCGCGA	CTGGCCGGGG	GGGGAGCGAT	GGCTTCGGTG	GCATTGCCGG
	64501	CCGGTGAGGT	CGGTCTGGTC	GAGGGCGTGT	GGATCGCGGC	CGGTAACGGC	CCCGCCTCGA
	64561	CAGTCGTGGC	CGGCGAGCCG	TCGGCGGTGG	AGGACGTGGT	GACGCGGTAT	GAGACCGAAG
	64621	GCGTGCGAGT	GCGTCGTATC	GCCGTCGACT	ACGCCTCCCA	CACGCCCCAC	GTGGAAACCA
5	64681	TCGAGGACGA	ACTCGCTGAG	GTACTGAAGG	GAGTTGCAGG	GAAGGCCGCG	TCGGTGGCGT
	64741	GGTGGTCGAC	CGTGGACAGC	GCCTGGGTGA	CCGAGCCGGT	GGATGAGAGT	TACTGGTACC
	64801	GGAACCTGCG	TCGCCCCGTC	GCGCTGGACG	CGGCGGTGGC	GGAGCTGGAC	GGGTCCGTGT
	64861	TCGTGGAGTG	CAGCGCCCAT	CCGGTGCTGC	TGCCGGCGAT	GGAACAGGCC	CACACGGTGG
	64921	CGTCGTTCGG	CACCGGTGAC	GGCGGTGGG	AGCGATGGCT	GACGGCGTTG	GCGCAGGCGT
10	64981	GGACCCCTGGG	CGCGGCAGTG	GACTGGGACA	CGGTGGTCGA	ACCGGTGCCA	GGCGGCTGTC
	65041	TCGATCTGCC	CACCTACGCG	TTCGAGCGCC	GGCGCTACTG	GCTGGAAGCG	GCCGGTGCCA
	65101	CCGACCTGTC	CGCGGCCGGG	CTGACAGGGG	CAGCACATCC	CATGCTGGCC	GCCATCACGG
	65161	CACTACCCGC	CGACGACGGT	GGTGTGTTC	TCACCGGCCG	GATCTCGTTG	CGCACGCATC
	65221	CTGGCTGGC	TGATCACGCG	GTGCGGGGCA	CGGTCTTGCT	GCCGGGCACG	GCCTTTGTGG
15	65281	AGCTGGTCAT	CCGGGCCGGT	GACGAGACCG	GTTGCGGGAT	AGTGGATGAA	CTGGTCATCG
	65341	AATCCCCCCT	CGTGGTGCCG	GCGACCGCAG	CCGTGGATCT	GTCGGTGACC	GTGGAAGGAG
	65401	CTGACGAGGC	CGGACGGCGG	CGAGTGACCG	TCCACGCCCC	CACCGAAGGC	ACCGGCAGCT
	65461	GGACCCGGCA	CGCCAGCGGC	ACCCTGACCC	CCGACACCCC	CGACACCCCC	AACGCTTCCG
	65521	GTGTTGTCCG	TGCGGAGCCG	TTCTCGCAGT	GGCCACCTGC	CACTGCCGCG	GCCGTCGACA
20	65581	CCTCGGAGTT	CTACTTGCGC	CTGGACGCGC	TGGGCTACCG	GTTCCGGACCC	ATGTTCCGCG
	65641	GAATGCGGGC	TGCTTGGCGT	GATGGTGACA	CCGTGTACGC	CGAGGTCCGCG	CTCCCCGAGG
	65701	ACCGTGCCGC	CGACGCGGAC	GGTTTCGGCA	TGCACCCGGC	GCTGCTCGAC	GCGGCTTTGC
	65761	AGAGCGGCAG	CCTGCTCATG	CTGGAATCGG	ACGGCGAGCA	GAGCGTGCAA	CTGCCCTTCT
	65821	CCTGGCACGG	CGTCCGGTTC	CACGCGACGG	GCGCGACCAT	GCTGCGGGTG	GCGGTGCTAC
25	65881	CGGGCCCGGA	CGGCCTCCGG	CTGCATGCCG	CGGACAGCGG	GAACCGTCCC	GTCGCGACGA
	65941	TCGACGCGCT	CGTGACCCGG	TCCCCGGAAG	CGGACCTCGC	GCCCGCCGAT	CCGATGCTGC
	66001	GGGTCGGGTG	GGCCCCGGTG	CCCGTACCTG	CCGGGGCCGG	TCCGTCCGAC	GCGGACGTGC
	66061	TGACGCTGCG	CGGCGACGAC	GCCGACCCGC	TCGGGGAGAG	CCGGGACCTG	ACCACCCGTG
	66121	TTCTCGACGC	GCTGCTCCGG	GCCGACCCGC	CGGTGATCTT	CCAGGTGACC	GGTGGCCTCG
30	66181	CCGCCAAGGC	GGCCGCAAGC	CTGGTCCGCA	CCGCTCAGAA	CGGACAGCCC	GGCCGCTTCT
	66241	TCCTCGTCGA	AACGGACCCG	GGAGAGGTCC	TGGACGGCGC	GAAGCGCGAC	GCGATCGCGG
	66301	CACTCGGCGA	GCCCCATGTG	CGGCTGCGCG	ACGGCCTCTT	CGAGGCAGCC	CGGCTGATGC
	66361	GGGCCACGCC	GTCCCTGACG	CTCCCGGACA	CCGGGTCTGT	GCAGCTGCGG	CCGTCCGCCA
	66421	CCGGTTCCCT	CGACGACCTT	GCCGTCGTCC	CCACCGACGC	CCCGGACCGG	CCGCTCGCGG
35	66481	CCGGCGAGGT	GCGGATCGCG	GTACGCGCGG	GGGGCCTGAA	CTTCCGGGAT	GTCACGGTCC
	66541	CGCTCGGTGT	GGTCGCGGAT	GCGCGTCCCG	TCGGCAGCGA	GGCCCGGGGT	GTCGTCTTGG
	66601	AGACCGGCCC	CGGTGTGCAC	GACCTGGCGC	CCGGCGACCG	GGTCTGGGGG	ATGCTCGCGG
	66661	GCGCCTTCGG	ACCGGTTCGCG	ATCACCGACC	GGCGGCTGCT	CGGCCGGATG	CCGGACGGCT
	66721	GGACGTTCCT	GCAGGCGGCG	TCCGTGATGA	CCGCGTTCCG	GACCGCGTGG	TACGGCCTGG
40	66781	TCGACCTGGC	CGGGCTGCGC	CCCGGCGAGA	AGGTCTTGAT	CCACGCGGCG	GCGACCGGTG
	66841	TCGGCGCGGC	GGCCGTCCAG	ATCGCGCGGC	ATCTGGGCGC	GGAGGTGTAC	GCGACACCA
	66901	GCGCCGCGAA	GCGCCATCTG	GTGGACCTGG	ACGGAGCGCA	TCTGCCCGAT	TCCCGCAGUA
	66961	CCGCGTTCGC	CGACGCGTTC	CCGCCGCTCG	ATGTCGTGCT	CAACTCGCTC	ACCGGTGAAT
	67021	TCCTCGACGC	GTCCGTTCGGC	CTGCTCGCGG	CGGGTGGCCG	GTTTCATCGAG	ATGGGGAAGA
45	67081	CGGACATCCG	GCACGCCGTC	CAGCAGCCGT	TCGACCTGAT	GGACGCCGGC	CCCGACCGGA
	67141	TGACGCGGAT	CATCGTCGAG	CTGCTCGGCC	TGTTCCGCGC	CGACGTGCTG	CACCCGCTGC
	67201	CGGTCCACGC	CTGGGACGTG	CGGCAGGCGC	GGGAGGCGTT	CGGCTGGATG	AGCAGCGGGC
	67261	GTCACACCGG	CAAGCTGGTG	CTGACGGTCC	CGCGGCCGCT	GGATCCCGAG	GGGGCCGTCG
	67321	TCATCACCGG	CGGCTCCGGC	ACCCTCGCCG	GCATCCTCGC	CCGCCACCTG	GGCCACCCCC
50	67381	ACACCTACCT	GCTCTCCCGC	ACCCACCCCC	CCGACACCAC	CCCCGGCACC	CACCTCCCTT
	67441	GCGACGTCGG	CGACCCCCAC	CAACTCGCCA	CCACCTCGC	CCGCATCCCC	CAACCCCTCA
	67501	CCGCCGTCTT	CCACACCGCC	GGAACCCCTG	ACGACGCCCT	GCTCGACAAC	CTCACCCCGG
	67561	ACCGCGTCGA	CACCGTCCCT	AAACCCAAAG	CCGACGCCGC	CTGGCACCTG	CACCGGCTCA
	67621	CCCGCGACAC	CGACCTCGCC	GCGTTCGTGC	TCTACTCCGC	GGTCGCCCGC	CTCATGGGCA
55	67681	GCCCGGGGCA	GGGCAACTAC	GTCCGCGGCA	ACGCGTTCCT	CGACGCGCTC	GCCGAACACC
	67741	GCCGTGCGCA	AGGGCTGCCG	GCGCAAGTCCC	TCGTCATGGG	CATGTGGGCG	GACGTACGCG
	67801	CGCTCACCGC	GAAACTCACC	GACGCGGACC	GCCAGCGCAT	CCGGCGCAGC	GGATTCCCGC
	67861	CGTTGAGCGC	CGCGGACGGC	ATGCGGCTGT	TCGACGCGGC	GACGCGTACC	CCGGAACCGG
	67921	TCGTCTGCGC	GACGACCGTC	GACCTCACCC	AGCTCGACGG	CGCCGTGCGC	CCGTTGCTCC
60	67981	GCGGTCTGGC	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CGCCCCGAAC	GCCGGCGAAG
	68041	AGCCCCTGCG	CGTGCGTCTT	GCCGGGCGTA	CCGCCGCCGA	GCAGCGGCGC	ATCATGACGG
	68101	AGGTCGTGCT	CCGCCACGCG	GCCGCGGTCC	TCGCGTACGG	GCTGGGCGAC	GCGGTGCGCG
	68161	CGGACGCTCC	GTTCCGCGAG	CTCGGTTTCG	ATTGCTGAC	CGCGGTGCGC	CTGCGCAATC
	68221	GGCTCGCGGC	CGAGACGGGG	CTGCGGCTGC	CGACGACGCT	GGTGTTCAGC	CACCCGACGG



	68281	CGGAGGCGCT	CACCGCCAC	CTGCTCGACC	TGATCGACGC	TCCCACCGCC	CGGATCGCCG
	68341	GGGAGTCCCT	GCCCGCGGTG	ACGGCCGCTC	CCGTGGCGGC	CGCGCGGGAC	CAGGACGAGC
	68401	CGATCGCCAT	CGTGGCGATG	GCGTGCCGGC	TGCCCCGGTG	TGTGACGTCG	CCCGAGGACC
5	68461	TGTGGCGGCT	CGTCGAGTCC	GGCACCAGACG	CGATCACCAC	GCCTCCTGAC	GACCGCGGCT
	68521	GGGACGTCGA	CGCGCTGTAC	GACGCGGACC	CGGACGCGGC	CGGCAAGGCG	TACAACCTGC
	68581	GGGGCGG'FTA	CCTGGCCGGG	GCGGCGGAGT	TCGACGCGGC	GTTCTTCGAC	ATCAGTCCGC
	68641	GCGAAGCGCT	CGGCATGGAC	CCGCAGCAAC	GCCTGCTGCT	CGAAACGGCG	TGGGAGGCGA
	68701	TCGAGCGCGG	CCGGATCAGT	CCGGCGTCGC	TCCGCGGCCG	GGAGGTCGGC	GTCTATGTCG
10	68761	GTGCGGCCGC	GCAGGGCTAC	GGGCTGGGCG	CCGAGGACAC	CGAGGGCCAC	GCGATCACC
	68821	GTGGTTCCAC	GAGCCTGCTG	TCCGGACGGC	TGGCGTACGT	GCTCGGGCTG	GAGGGCCCGG
	68881	CGGTACCCGT	GGACACGGCG	TGCTCGTCTG	CTCTGGTCGC	GCTGCATCTG	GCGTGCCAGG
	68941	GGGTACGCTT	GGGCGAGTGC	GAATCGCTC	TGGCCGGAGG	GGTCTCCGTA	CTGAGTTCGC
	69001	CGGCCGCGTT	CGTGGAGTTC	TCCCGCCAGC	GCGGGCTCGC	GGCCGCAAGT	ATCGGCTAAGT
15	69061	CGTTCGGCGC	GGGCGCGGAC	GGCAGCAGCT	GGTCCGAGGG	CGTGGGCGTG	CTCGTACTGG
	69121	AACGGCTCTC	CGACGCCGAG	CGGCTCGGGC	ACACCGTGCT	CGCCGTCGTC	CGCGGCAGCG
	69181	CCGTACCGTC	CGACGGCGCC	TCCAACGGCC	TCACCGCGCC	GAACGGGCTC	TCGCAGCAGC
	69241	GGGTATCCCG	GAAGGCGCTC	GCCGCGGCCG	GGCTGACCGG	CGCCGACGTG	GACGTCGTCG
	69301	AGGGGACCGG	CACCGGCACC	CGGCTCGGGC	ACCCGGTCGA	GGCGGACGCG	CTGCTCGCGA
20	69361	CGTACGGGCA	GGACCGTCCG	GCACCGGCTC	GGCTGGGCTC	GCTGAAGTCG	AACATCGGAC
	69421	ATGCCACGGC	CGCGGCCGGT	GTCGCGGGCG	TCATCAAGAT	GGTGCAGGCG	ATCGGCGCGG
	69481	GCACGATGCC	GCGGACGCTG	CATGTGGAGG	AGCCCTCGCC	CGCCGTCGAC	TGGAGCACC
	69541	GACAGGTGTC	CCTGCTCGGC	TCCAACGGCC	CCTGGCCGGA	CGACGAGCGT	CCGCGCCGGG
	69601	CGGCCGTCTC	CGCGTTCGGG	CTCAGCGGGA	CGAACGCGCA	CGTCATCCTG	GAACAGCACU
25	69661	GTCCGGCGCC	CGTGGCGTCC	CAGCCGCCCC	GGCCGCCCCG	TGAGGAGTCC	CAGCCGCTGC
	69721	CGTGGGTGCT	CTCCGCGCGG	ACTCCGGCCG	CGTGGCGGGC	CCAGGCGGGC	CGGCTGCGCG
	69781	ACCACTCGC	GGCGGCACCG	GACGCGGATC	CGTTGGACAT	CGGGTACGCG	CTGGCCACCA
	69841	GCCGCGCCCA	GTTGCGCCAC	CGTGCCGCGG	TCGTGCGCCAC	CACCCCGGAC	GGATTCCGTG
	69901	CCGCGCTCGA	CGGCCTCGCG	GACGGCGCGG	AGGCGCCCGG	AGTCGTCACC	GGGACCCTC
30	69961	AGGAGCGGGC	CGTCGCCTTC	CTCTTCGACG	GCCAGGGCGC	CCAGCGCGCC	GGAATGGGGC
	70021	GCGAGCTCCA	CCGCCGGTTC	CCCGTCTTCG	CCGCCGCGTG	GGACGAGGTC	TCCGACGCGT
	70081	TCGGCAAGCA	CCTCAAGCAC	TCCCCACGG	ACGTCTACCA	CGGCGAACAC	GGCGCTCTCG
	70141	CCCATGACAC	CCTGTACGCC	CAGGCCGGCC	TGTTACGCT	CGAAGTGGCG	CTGCTGCGGC
	70201	TGCTGGAGCA	CTGGGGGGTG	CGGCCGACG	TGCTCGTCGG	GCATCCGTC	CTGCTGGCTGA
35	70261	CCGCGGCGTA	CGCGGCGGGG	GTGCTACCC	TGGCGGACGC	GACGGAGTTG	ATCGTGGCCC
	70321	GGGGGCGGGC	GCTGCGGGCG	CTGCCGCCCG	GGGCGATGCT	CGCCGTCGAC	GGAAGCCCGG
	70381	CGGAGGTCGG	CGCCCGCACG	GATCTGGACA	TCGCCGCGGT	CAACGGCCCG	TCCGCCGTGG
	70441	TGCTCGCCGG	TTCGCCGGAC	GATGTGGCGG	CGTTCGAACG	GGAGTGGTCG	GCGGCCGGGC
	70501	GGCGCACGAA	ACGGCTCGAC	GTCGGGCACG	CGTTCCACTC	CCGGCACGTC	GACGGTGCGC
40	70561	TCGACGGCTT	CCGTACGGTG	CTGGAGTCGC	TCGCGTTCGG	CGCGGCGCGG	CTGCCGCTGG
	70621	TGTCCACGAC	GACGGGCGCG	GACGCGCGG	ACGACCTCAT	AACGCCCGCG	CATGGCTGTC
	70681	GCCATGCGCG	TCGGCCGGTG	CTGTTCTCGG	ATGCCGTCCG	GGAGCTGGCC	GACCGCGGCG
	70741	TCACCACGTT	CGTGGCCGTC	GGCCCCCTCC	GCTCCCTGGC	GTGCGCCGCG	GCGGAGAGCG
	70801	CCGGGGAGGA	CGCCGGGACC	TACCACGCGG	TGCTGCGCGC	CCGGACCGGT	GAGGAGACCG
45	70861	CGGCGCTGAC	CGCCCTCGCC	GAGCTGCACG	CCCACGGCGT	CCCGGTTCGAC	CTGGCCGCGG
	70921	TACTGGCCGG	TGGCCGGGCA	GTGGACCTTC	CCGTGTACGC	GTTCCAGCAC	GTTCTCTACT
	70981	GGCTGGCCCC	GGCCGTGGCG	GGGGGCGCGG	CCACCGTGGC	GGACCGCGGG	CGTCCGCGCG
	71041	AGTCCGAGCC	GGAGGACCTC	ACCGTCGCGG	AGATCGTCCG	TCGGCGCACC	GCGGCGCTGC
	71101	TCGGCGTCAC	GGACCCCGCC	GACGTCGATG	CGGAAGCGAC	GTTCTTCGCG	CTCGGTTTCG
50	71161	ACTCACTGGC	GGTGCAGCGG	CTGCGCAACC	AGCTCGCCTC	GGCAACCGGG	CTGGACCTGC
	71221	CGGCGGCCGT	CCTGTTTCGAC	CACGACACCC	CGGCCGCGCT	CACCGCGTTC	CTCCAGGACC
	71281	GGATCGAGGC	CGGCCAGGAC	CGGATCGAGG	CCGGCGAGGA	CGACGACGCG	CCCACCGTGC
	71341	TCTCGCTCCT	GGAGGAGATG	GAGTCGCTCG	ACGCCGCGGA	CATCGCGGCG	ACGCCGGCCC
	71401	CGGAGCGTGC	GGCCATCGCC	GATCTGCTCG	ACAAGCTCGC	CCATACCTGG	AAGGACTACC
55	71461	GATGAGCACC	GATACGCACG	AGGGAACGCC	GCCCCCGGCG	CGCTGCCCAT	TCGCGATCCA
	71521	GGACGGTCAC	CGCGCCATCC	TGGAGAGCGG	CACGGTGGGT	TCGTTTCGACC	TGTTTCGGCGT
	71581	CAAGCACTGG	CTGGTCGCCG	CCGCCGAGGA	CGTCAAGCTG	GTCACCAACG	ATCCGCGGTT
	71641	CAGCTCGGCC	GCGCCGTCCG	AGATGCTGCC	CGACCGGCGG	CCCGGCTGGT	TCTCCGGGAT
	71701	GGACTCACCG	GAGCACAACC	GCTACCGGCA	GAAGATCGCG	GGGGACTTCA	CACTGCGCGC
60	71761	GGCGCGCAAG	CGGAGGACT	TCGTGCGCCA	GGCCGCGGAC	GCCTGCCTGG	ACGACATCGA
	71821	GGCCGCGGGA	CCCGGCACCG	ACCTCATCCC	CGGGTACGCC	AAGCGGCTGC	CCTCCCTCGT
	71881	CATCAACGCG	CTGTACGGGC	TCACCCCTGA	GGAGGGGGCC	GTGCTGGAGG	CACGGATGCG
	71941	CGACATCACC	GGCTCGGCCG	ATCTGGACAG	CGTCAAGACG	CTGACCGACG	ACTTCTTCGG
	72001	GCACGCGCTG	CGGCTGGTCC	GCGCGAAGCG	TGACGAGCGG	GGCGAGGACC	TGCTGCACCG
	72061	GCTGGCCTCG	GCCGACGACG	GCGAGATCTC	GCTCAGCGAC	GACGAGGCGA	CGGGCGTGTT



	72121	CGCGACGCTG	CTGTTCGCCG	GCCACGACTC	GGTGCAGCAG	ATGGTCGGCT	ACTGCCTCTA
	72181	CGCACTGCTC	AGCCACCCCG	AGCAGCAGGC	GGCGCTGCGC	GCGCGCCCGG	AGCTGGTCTGA
	72241	CAACGCGGTC	GAGGAGATGC	TCCGTTTCCT	GCCCGTCAAC	CAGATGGGCG	TACCGCGCGT
	72301	CTGTGTCGAG	GACGTGCTATG	TGCGGGGCGT	GCGCATCCGT	GCGGGCGACA	ACGTGATCCC
5	72361	GCTCTACTCG	ACGGCCAACC	GCGACCCCGA	GGTGTTCCTG	CAGCCCGACA	CCTTCGATGT
	72421	GACGCGCCCG	CTGGAGGGCA	ACTTCGCGTT	CGGCCACGGC	ATTACAAGT	GTCCCGGCCA
	72481	GCACATCGCC	CGGGTGCTCA	TCAAGGTCGC	CTGCCTGCGG	TTGTTTCGAGC	GTTTCCCGGA
	72541	CGTCCGGCTG	GCCGGCGACG	TGCCGATGAA	CGAGGGGCTC	GGGCTGTTCA	GCCCGGCCGA
10	72601	GCTCGGGGTC	ACCTGGGGGG	CGGCATGAGT	CACCCGGTGG	AGACGTTGCG	GTTGCCGAAC
	72661	GGGACGACGG	TCGCGCACAT	CAACGCGGGC	GAGGCGCAGT	TCCTCTACCG	GGAGATCTTC
	72721	ACCCAGCGCT	GCTACCTGCG	CCACGGTGTC	GACCTGCGCC	CGGGGGACGT	GGTGTTCGAC
	72781	GTCGGCGCGA	ACATCGGCAT	GTTACGCTT	TTCGCGCATC	TGGAGTGTCC	TGGTGTGACC
	72841	GTGCACGCCT	TCGAGCCCCG	GCCCGTGCCG	TTCGCGGCGC	TGCGGGCGAA	CGTGACGCGG
	72901	CACGGCATCC	CGGGCCAGGC	GGACAGTGTC	GCGGTCTCCG	ACAGCTCCGG	CACCCGGAAG
15	72961	ATGACCTTCT	ATCCCGACGC	CACGCTGATG	TCCGGTTTCC	ACGCGGATGC	CGCGGCCCGG
	73021	ACGGAGCTGT	TGCGCAGCTG	CGGCCTCAAC	CGCGGCTACA	CCGCGGAGGA	CGTCGACACC
	73081	ATGCTCGCGC	AACTGCCCCA	CGTCAGCGAG	GAGATCGAAA	CCCCTGTGGT	CCGGCTCTCC
	73141	GACGTCAATC	CGGAGCGCGG	TATCGAGGCC	ATCGGCCTGC	TGAAGGTCGA	CGTGGAGAAG
	73201	AGCGAACGGC	AGGTCTTCGC	CGGCCTCGAG	GACACCGACT	GGCCCCGTAT	CCGCCAGGTC
20	73261	GTCGCGGAGG	TCCACGACAT	CGACGGCGCG	CTCGAGGAGG	TCGTACACGT	GCTCCGCGGC
	73321	CATGGCTTCA	CCGTGGTTCG	CGAGCAGGAA	CCGCTGTTCG	CCGGCACGGG	CATCCACCAG
	73381	GTCGCCGCGC	GGCGGGTGGC	CGGCTGAGCG	CCGTCGGGGC	CGCGGCCGTC	CGCACCGGCG
	73441	GCCGCGGTGC	GGACGGCGGC	TCAGCCGGCG	TCGGACAGTT	CCTTGGGCGA	TTGTGACCG
	73501	CCCTTCAACC	CCAGCTTGCG	GAACACGTTG	GTGAGGTGCT	GTTCCACCGT	GTTGGAGGTC
25	73561	ACGAACAGCT	GGCTGGCGAT	CTCCTTGTTG	GTGCGCCCGA	CCGCGGCGTG	CGACGCCACC
	73621	CGCCGCTCCG	CCTCGGTCAG	CGATGTGATC	CGCTGCGCCG	GCGTCACGTC	CTGGGTGCCG
	73681	TCCGCGTCCG	AGGACTCCCC	ACCGAGCCGC	CGGAGGAGCG	GCACGGCTCC	GCACTGGGTC
	73741	GCGAGGTGCC	GTGCGCGGCG	GAACAGTCCC	CGCGCACGGC	TGTGCCGCGG	GAGCATGCCG
	73801	CACGCTTCGC	CCATGTTCGC	GAGGACGCGG	CCCAGCTCGT	ACTGGTCGCG	GCACTGATG
30	73861	AGCAGATCGG	CGGCCTCGTC	GAGCAGTTTC	ATCCGCTTGG	CCGCGGCACT	GTAGGCCGCC
	73921	TGCACCCGCA	GCGTCATCAC	CCGCGCCCGG	GACCCCATCG	GCCGGGACAG	CTGCTCGGAG
	73981	ATGAGCCTCA	GCCCCCTCGT	ACGGCCGCGG	CCGAGCAGCA	GAAGCGCTTC	GGCGGCGTCG
	74041	ACCCGCCACA	GGGCCAGGCC	CGGCACGTCG	ACGGACCAGC	GTGCGATCCG	CTCCCCGCAG
	74101	TCCCGGAACG	CGTTGTACGC	CGCCCGGTAC	CGCCCGGCCG	CGAGATGGTG	TTGCCACCGG
35	74161	GCCCAGACCA	TGTGCAGTTC	GAAGAGGCTG	TCGGAGGTCT	CCTCCGGCAA	CGGCTCGGCG
	74221	AGCCACCGCT	CCGCCCGGTC	CAGGTGCGCC	AGTCGGATCG	CGGCGGCCAC	GGTGTGCTC
	74281	AGCGGCAATG	CGGCGGCCAT	CCCCCAGGAG	GGCACGACCC	GGGGGGCGAG	CGCGGCCTCG
	74341	CCGCATTCTGA	CGGCGGCGGT	CAGGTGCGCG	CGGCGCAGCG	CGGCCTCGGC	GCGGAACCCC
	74401	GCGTGGACCG	CCTCGTCCGC	CGGGGTCCGC	ATGTTGTCTG	CACCGGCCAG	CTTGTGACCC
40	74461	CAGGACTGGA	CGGCATCGGT	GTCCTCGGCG	TAGAGCAGGG	CCAGCAACGC	CATCATGGTC
	74521	GTGGTCCGGT	CCGTCTGTGAC	CCGGGAGTGC	TGGAGCACGT	ACTCGGCTTT	GGCCTCGGCC
	74581	TGTTCCGACC	AGCCGCGCAG	CGCCTTGCTC	AGGGCCTTGT	CGGCGACGGC	CGGCTCGCCG
	74641	ACGGTCCCGG	AAAACGAGGC	GACCTGCTCC	TCGGCCGGCG	GATCGGCGCG	ACGCGGCGGA
	74701	TCGGCCGCGC	CGGGATAGAT	CAGCGCGAGG	GACAGGTCCG	CGACGCGCAG	GTGCGCCCGG
45	74761	CCCTGCTCGC	TCGGGGCGGC	GGAGCGCTGG	GCCGCCAGGA	CCTCGGCGGC	CTCGCCCGGC
	74821	CGCCCGTCCA	TCGCCAGCCA	GCAGGCGAGC	GACACGGCGT	GCTCGCTGGA	GAGGAGCCGT
	74881	TCCCGCGACG	CGGTGAGCAG	CTCGGGCACA	TGCCGGCCCG	ATCTGGCGGG	ATCGCAGAGC
	74941	CGCTCGATGG	CGGCGGTGTC	GACGCGCAGT	GCGGCGTGGA	CGGCGGGGTC	GTGCGAGGCC
50	75001	CGGTAGGCGA	ACTCCAGGTA	GGTGACGGCC	TCGTGAGGCT	CGCCGCGCAG	GTAGGTGCTC
	75061	CGCGCGGCGT	CGGTGAACAG	CCCCGCGACC	TCGGCGCCGT	GCACCCGGCC	GGTACCCATC
	75121	TGGTGGCGGG	CGAGCACCTT	GCTGGCCACG	CCGCGGTCCC	GCAGCAGTTC	CAGCGCCAGC
	75181	TCGTGACGGC	CACGCGGCTC	GGCGGCGGAG	AGGTGCTCGA	GTACGACGGA	GCGGGCCGCG
	75241	GGGTGCGGGA	ACCGCCCTTC	CCGCGACAGC	CGCCCCCTCG	CCAGCTGTTT	GTGGGCCTGC
	75301	TCGACCGCCT	CGGTGTTCGAG	GCCGGTCATC	CGCTGGACGA	GGGTGAGTTC	GACACTCTCG
55	75361	CCGAGCACGG	CGGAAGCTCG	GGCGACGCTC	AGCGCGGCGG	GGCCGCAACG	ATAGAGCGAC
	75421	CCGAGGTAGG	CGAGCCGGTA	GCGCCGCCCC	GCGACCACTT	CCAGGCAACC	TGAGGTCCGT
	75481	GTCCGTGCCT	CCCGGATGTC	GTCGATCAGG	CCGTGGCCGA	GGAGCAGGTT	GCCGCGGGTC
	75541	GCCCGGAACG	CCTGGGCCAC	CACGTCTGTC	TGCGCGTCCT	GGCCGAGGTG	CCGGCGCACG
	75601	AGTTCCGGTG	TCTGCGCCTC	GGTGAGCGGG	CGCAGCGCGA	TCTCTGGTGA	GTGGCGCAGA
60	75661	CTCAGCAGTG	CCGCCCCGAA	TTGGGAGTGG	GCGGGCGTCG	GCCGGAGCAG	CTCGGTCAGC
	75721	ACGATGGCGA	CACGGGCCCC	GCTGATGCGG	CGCGCGAGGT	GGAGCAGGCA	GCGCAGCGAC
	75781	GGCGCGTCCG	CGTGGTGCAC	GTCGTCTGATG	CCGATCAGTA	CGGGCCGCTC	CGCGCGGAGC
	75841	GTCAGCACCG	TGCGGGTGAG	TTCGTTCCCC	AGGCGGTTGT	CGACGTCGCG	CGCAGGTTT
	75901	TCGCAAGATG	CCGTCAAGCG	GACCAGCTCC	GGTGTCCGGG	CGGCCAGCTC	GGGCTGGTCC

75961 AGGAGCTGGC CGAGCATGCC GTACGGCAGG GCCCGCTCCT CCATGGAGCA CACCCGCGCA  
 76021 AGGGTGACGA AGCCGGCCTT GGCCGCGGCG GCGTCGAGGA GTTCGGTCTT GCCGCAGGCG  
 76081 ATCGGCCCCG TGACGGCGGC GACGACGCC CGCCCGCCCC CCGCTCGGGT GAGCGCCCGG  
 76141 TGGAGGGAAC CGAACTCGTC ATCGCGGGCG ATCAGGTCTG GGGGAGATAA GCGCGCTATC  
 5 76201 ACGAATGGAA CTACCTCGCG ACCGTCGTGG AAACCCATAG GCATCACATG GCTTGTGAT  
 76261 CTGTACGGCT GTGATTCAGC CTGGCGGGAT GCTGTGCTAC AGATGGGAAG ATGTGA<sup>2</sup>CTA  
 76321 GGGCCGTGCC GTTCCTCAG GAGCCGACCG CCCCCGGCGC CACCCGCCGT ACCCCCTGGG  
 76381 CCACCAGCTC GGCGACCCGC TCCTGGTGGT CGACGAGGTA GAAGTGCCCG CCGGGGAAGA  
 76441 CCTCCACCGT GGTCGGCGCG GTCGTGTGCC CGGCCAGGC GTGGGCCTGC TCCACCGTCG  
 10 76501 TCTTCGGATC GTCGTCACCG ATGCACACCG TGATCGGCGT CTCCAGCGGC GCGCGGGCT  
 76561 CCCACCGGTA CGTCTCCGCC GCGTAGTAGT CCGCCCGCAA CGGCGCCAGG ATCAGCGCGC  
 76621 GCATTTTCGTC GTCCGCCATC ACATCGGCGC TCGTCCCGCC GAGGCCGATG ACCGCCGCCA  
 76681 GCAGCTCGTC GTCGGACGCG AGGTGGTCTT GGTCGGCGCG CGGCTGCGAC GCGCCCCGCC  
 76741 GGCCCGAGAC GATCAGGTGC GCCACCGGGA GCCGCTGGG CAGCTCGAAC GCGAGTGTG  
 15 76801 CGCCCATGCT GTGGCGGAAC AGCACCAGCG GACGGTCCAG CCCCGGCTTC AACGCTCGG  
 76861 CCACGAGGCC GGCGAGAACA CGCAGGTCGC GCACCGCTC CTCGTCGCGG CGGTCTGGC  
 76921 GGCCGGGGTA CTGCACGGCG TACACGTCCG CCACCGGGG GAGCGCACGG GCCAGCGGAA  
 76981 GGTAGAACGT CGCCGATCCG CCGGCGTGGG GCAGCAGCAC CACCCGTACC GGGGCCTCGG  
 77041 GCGTGGGGAA GAACTGCCGC AGCCAGAGTT CCGAGCTCAC CGCACCCCT CCGCCGCGAC  
 20 77101 CTGGGGAGCC CGGAACCGG TGATCTCGCG CAAGTGCTTC TCCCGCATCT CCGGGTCGGT  
 77161 CACGCCCCAT CCTCCTCCG GCGCCAGACA GAGGACCCG ACTTTGCCGT TGTGCACATT  
 77221 GCGATGCACA TCGCGACCG CCGACCCGAC GTCGTCGAGC GGTAGGTC ACGCAGCGT  
 77281 CGGGTGCACC ATCCCTTGC AGATCAGGCG GTTCGCCTC CACGCCTAC GATAGTTGCG  
 77341 GAAGTGGGTA CCGATGATCC GCTTCACGGA CATCCACAGG TACCGATTGT CAAAGGCGTG  
 25 77401 CTCGTATCCC GAGGTTGACG CGCAGGTGAC GATCGTGCCA CCCCACGTG TCACGT<sup>3</sup>GAC  
 77461 ACTCGCGCCG AACGTCGCGC GCCCCGGGTG CTCGAACACG ATGTGGGGAT CGTCACCGCC  
 77521 GGTCAGCTCC CGGATC

Those of skill in the art will recognize that, due to the degenerate nature of the  
 30 genetic code, a variety of DNA compounds differing in their nucleotide sequences can be  
 used to encode a given amino acid sequence of the invention. The native DNA sequence  
 encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to  
 illustrate a preferred embodiment of the invention, and the present invention includes  
 DNA compounds of any sequence that encode the amino acid sequences of the  
 35 polypeptides and proteins of the invention. In similar fashion, a polypeptide can typically  
 tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid  
 sequence without loss or significant loss of a desired activity. The present invention  
 includes such polypeptides with alternate amino acid sequences, and the amino acid  
 sequences shown merely illustrate preferred embodiments of the invention.

40 The recombinant nucleic acids, proteins, and peptides of the invention are many  
 and diverse. To facilitate an understanding of the invention and the diverse compounds  
 and methods provided thereby, the following general description of the FK-520 PKS  
 genes and modules of the PKS proteins encoded thereby is provided. This general  
 description is followed by a more detailed description of the various domains and  
 45 modules of the FK-520 PKS contained in and encoded by the compounds of the  
 invention. In this description, reference to a heterologous PKS refers to any PKS other  
 than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference

to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

5           The FK-520 PKS is composed of three proteins encoded by three genes designated *fkfA*, *fkfB*, and *fkfC*. The *fkfA* ORF encodes extender modules 7 - 10 of the PKS. The *fkfB* ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The *fkfC* ORF encodes extender modules 5 - 6 of the PKS. The *fkfP* ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520  
10 polyketide.

          The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihydroxycyclohexene carboxylic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the  
15 FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the  
20 heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin, FK-506, rifamycin, and avermectin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the  
25 coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

          In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA  
30 ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction with a coding sequence for a PKS that  
35 synthesizes FK-520, an FK-520 derivative, or another polyketide.

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA, a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded  
5 thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is  
10 merely added to coding sequences for modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

15 In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the  
20 DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such  
25 domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding  
30 sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

In an illustrative embodiment of this aspect of the invention, the invention  
35 provides recombinant PKSs and recombinant DNA compounds and vectors that encode

such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS  
5 encoded thereby can be fed or supplied with N-acylcysteamine thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

The second extender module of the FK-520 PKS includes a KS, an AT specific  
10 for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the  
15 coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second  
20 extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid  
25 module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of  
30 these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-  
35 520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding

domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

The fourth extender module of the FK-520 PKS includes a KS, an AT that binds ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In

one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fourth extender  
5 module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS  
10 that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA  
15 specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, a PKS for a polyketide other than FK-520, or from  
20 chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender module of the FK-520 PKS.

25 As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by  
30 those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of

the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzymes of the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl-21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence



can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

5 In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-functional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the  
10 expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-19 to C-20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment,  
15 the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that  
20 express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

25 The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes  
30 the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In  
35 another embodiment, a DNA compound comprising a sequence that encodes the sixth

extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

5 In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, 10 DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding 15 sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant 20 FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an 25 extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as, for example, the coding sequences for extender module two encoded by the *eryAI* gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant 30 activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of 35 the FK-506 PKS. The invention also provides recombinant host cells derived from FK-

506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have  
5 been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of  
10 applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding  
15 sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

20 In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting the KR, the DH, and/or the ER; and/or replacing the  
25 KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous  
30 seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes  
5 code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an  
10 illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-  
15 506 but that express this recombinant PKS and so synthesize the corresponding (C-15-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

In another illustrative embodiment, the present invention provides a hybrid PKS  
20 in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

25 The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520  
30 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another  
35 embodiment, a DNA compound comprising a sequence that encodes the eighth extender

module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, and an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence

for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a  
5 DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the tenth extender module coding  
10 sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP.  
15 In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that  
20 synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The  
25 enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the *fkbP* gene and so provides recombinant methods for expressing the *fkbP* gene product in recombinant host cells. The recombinant *fkbP* genes of the invention include those in which the  
30 coding sequence for the adenylation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of  
35 pipecolic acid (see Nielsen *et al.*, 1991, *Biochem.* 30: 5789-96). The *fkbL* gene encodes a

homolog of RapL, a lysine cyclodeaminase responsible in part for producing the  
pipecolate unit added to the end of the polyketide chain. The *fkbb* and *fkbl* recombinant  
genes of the invention can be used in heterologous hosts to produce compounds such as  
FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel  
5 polyketides and non-ribosomal peptides.

The present invention also provides recombinant DNA compounds that encode  
the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520.  
Figure 2 shows the various sites on the FK-520 polyketide core structure at which these  
enzymes act. By providing these genes in recombinant form, the present invention  
10 provides recombinant host cells that can produce FK-520. This is accomplished by  
introducing the recombinant PKS, P450 oxidase, and methyltransferase genes into a  
heterologous host cell. In a preferred embodiment, the heterologous host cell is  
*Streptomyces coelicolor* CH999 or *Streptomyces lividans* K4-114, as described in U.S.  
Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar.  
15 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by  
reference. In addition, by providing recombinant host cells that express only a subset of  
these genes, the present invention provides methods for making FK-520 precursor  
compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds  
20 and vectors that are useful in generating, by homologous recombination, recombinant  
host cells that produce FK-520 precursor compounds. In this aspect of the invention, a  
native host cell that produces FK-520 is transformed with a vector (such as an SCP2\*  
derived vector for *Streptomyces* host cells) that encodes one or more disrupted genes  
(i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those  
25 genes. When the vector integrates by homologous recombination, the native, functional  
gene is deleted or replaced by the non-functional recombinant gene, and the resulting  
host cell thus produces an FK-520 precursor. Such host cells can also be complemented  
by introduction of a modified form of the deleted or mutated non-functional gene to  
produce a novel compound.

30 In one important embodiment, the present invention provides a hybrid PKS and  
the corresponding recombinant DNA compounds that encode those hybrid PKS  
enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that  
comprises all or part of one or more modules and thioesterase/cyclase domain of a first  
PKS and all or part of one or more modules, loading module, and thioesterase/cyclase



domain of a second PKS. In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapamycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonyl CoA.

In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specific for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60/091,526, incorporated herein by reference. The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

(i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS, but also:

(ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally

occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,

- (iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520
- 5 PKS genes, and
- (iv) from combinations of the foregoing.

Various hybrid PKSs of the invention illustrating these various alternatives are described herein.

- Examples of the production of a hybrid PKS by co-expression of PKS genes from
- 10 the FK-520 PKS and another non-FK-520 PKS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples
- 15 include (i) replacement of the *fkfC* gene with the *rapB* gene; and (ii) replacement of the *fkfA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily
- 20 modified to act only as neurotrophins, as described in Example 6, below.

- Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the *fkfA* gene of an FK-520 or FK-506 producing host cell with a hybrid *fkfA* gene in which: (a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the
- 25 rapamycin PKS; and (b) the module 8 coding sequences have been replaced by the module 8 coding sequence of the rifamycin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13-desmethoxy-13-methyl-FK-520 or 13-
- 30 desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the producing host cell by a vector such as pHU204, which is a plasmid pRM5 derivative that has the well-characterized SCP2\* replicon, the *colE1* replicon, the *tsr* and *bla* resistance genes, and a *cos* site. This vector can be used to introduce the recombinant *fkfA* replacement gene in an FK-520 or FK-
- 35 506 producing host cell (or a host cell derived therefrom in which the endogenous *fkfA*

gene has either been rendered inactive by mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau *et al.*, 1999, "Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units," *Biochemistry* 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau *et al.*, *supra*. One can also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale *et al.*, 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," *Science* 284: 482-485, incorporated herein by reference.

The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the present invention:

#### **Avermectin**

U.S. Pat. No. 5,252,474 to Merck.

MacNeil *et al.*, 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

MacNeil *et al.*, 1992, *Gene* 115: 119-125, Complex Organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase.

Ikeda *et al.*, Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA* 96: 9509-9514.

**Candicidin (FR008)**

Hu *et al.*, 1994, *Mol. Microbiol.* 14: 163-172.

### **Epothilone**

U.S. Pat. App. Serial No. 60/130,560, filed 22 April 1999.

### **Erythromycin**

5 PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5,824,513 to Abbott.

Donadio *et al.*, 1991, *Science* 252:675-9.

Cortes *et al.*, 8 Nov. 1990, *Nature* 348:176-8, An unusually large  
multifunctional polypeptide in the erythromycin producing polyketide synthase of  
10 *Saccharopolyspora erythraea*.

### Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

### **FK-506**

Motamedi *et al.*, 1998, The biosynthetic gene cluster for the macrolactone ring of  
15 the immunosuppressant FK-506, *Eur. J. biochem.* 256: 528-534.

Motamedi *et al.*, 1997, Structural organization of a multifunctional polyketide  
synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur.*  
*J. Biochem.* 244: 74-80.

### Methyltransferase

20 US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from  
*Streptomyces* MA6858. 31-O-desmethyl-FK-506 methyltransferase.

Motamedi *et al.*, 1996, Characterization of methyltransferase and  
hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and  
FK-520, *J. Bacteriol.* 178: 5243-5248.

25 *Streptomyces hygroscopicus*

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

### **Lovastatin**

U.S. Pat. No. 5,744,350 to Merck.

### **Narbomycin**

30 U.S. patent application Serial No. 60/107,093, filed 5 Nov. 1998, and Serial No.  
60/120,254, filed 16 Feb. 1999.

### **Nemadectin**

MacNeil *et al.*, 1993, *supra*.

### **Niddamycin**

Kakavas *et al.*, 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

### **Oleandomycin**

Swan *et al.*, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano *et al.*, 1998, Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

### **Picromycin**

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue *et al.*, 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pikC*-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chemistry & Biology* 5(11): 661-667.

Xue *et al.*, Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity, *Proc. Natl. Acad. Sci. USA* 95: 12111-12116.

### **Platenolide**

EP Pat. App. Pub. No. 791,656 to Lilly.

### **Rapamycin**

Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA* 92:7839-7843.

Aparicio *et al.*, 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene* 169: 9-16.

### **Rifamycin**

August *et al.*, 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amiclatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

### **Sorangium PKS**

U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

### **Soraphen**

U.S. Pat. No. 5,716,849 to Novartis.

Schupp *et al.*, 1995, *J. Bacteriology* 177: 3673-3679. A *Sorangium cellulosum* (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

5     **Spiramycin**

U.S. Pat. No. 5,098,837 to Lilly.

Activator Gene

U.S. Pat. No. 5,514,544 to Lilly.

**Tylosin**

10       EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

Kuhstoss *et al.*, 1996, *Gene* 183:231-6., Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Tailoring enzymes

15       Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in  
20       constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the FK-520 PKS in PCT patent publication No. 98/51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073,538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

25       The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules  
30       one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived  
35       for example, from the picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08/989,332, filed 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cell/vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce  
5 actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.

The present invention provides a wide variety of expression vectors for use in  
10 *Streptomyces*. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2\* (see Hopwood *et al.*, *Genetic Manipulation of Streptomyces: A Laboratory manual* (The John Innes Foundation, Norwich, U.K., 1985); Lydiate *et al.*, 1985, *Gene* 35: 223-235; and Kieser and Melton, 1988, *Gene* 65: 83-91, each of which is incorporated herein by reference),  
15 SLP1.2 (Thompson *et al.*, 1982, *Gene* 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth *et al.*, 1989, *Mol. Gen. Genet.* 219: 341-348, and Bierman *et al.*, 1992, *Gene* 116: 43-49, each of which is incorporated herein by reference), or a high copy number vector, such as pIJ101 and pJV1 (see Katz *et al.*, 1983, *J. Gen. Microbiol.* 129: 2703-2714; Vara *et al.*, 1989, *J. Bacteriol.* 171: 5782-5781; and Servin-Gonzalez, 1993,  
20 *Plasmid* 30: 131-140, each of which is incorporated herein by reference). Generally, however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an *E. coli* origin of replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood  
25 *et al.*, *supra*).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin), *tsr* (confers resistance  
30 to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter,  
35 typically with an attendant ribosome binding site sequence. The present invention



provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the *fkfO* gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the *fkfO* and *fkfB* genes. The *fkfO* promoter is believed to be bi-directional in that it promotes transcription of the genes *fkfO*, *fkfP*, and *fkfA* in one direction and *fkfB*, *fkfC*, and *fkfL* in the other. Thus, in one aspect, the present invention provides a recombinant expression vector comprising the promoter of the *fkfO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkfO*. In a preferred embodiment the transcribed gene is an FK-520 PKS gene. In another preferred embodiment, the transcribed gene is a gene that encodes a protein comprised in a hybrid PKS.

Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the *actI* promoter and its attendant activator gene *actII-ORF4*, which is provided in the pRM1 and pRM5 expression vectors, *supra*. This promoter is activated in the stationary phase of growth when secondary metabolites are normally synthesized. Other useful *Streptomyces* promoters include without limitation those from the *ermE* gene and the *melC1* gene, which act constitutively, and the *tipA* gene and the *merA* gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to *Streptomyces* and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is inserted into a neutral site of the chromosome or in a vector under the control of the inducible *merA* promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the *actII-ORF4* gene discussed above include *dnrI*, *redD*, and *ptpA* genes (see U.S. patent application Serial No. 09/181,833, *supra*) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the

location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl CoA.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbl*, *fkbl*, and *fkbl* genes are sufficient to confer this ability on *Streptomyces* host cells. For conversion of 2-hydroxymalonyl to 2-methoxymalonyl, the *fkbl* gene is also employed. While the complete coding sequence for *fkbl* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence herein shows one T, there may be two, resulting in an extension of the *fkbl* reading frame to encode the amino acid sequence:

MTIVKCLVWDLNLTWRGTVLEDDEVVLTDREIVITLDDRGLQAVASKNDH  
DLAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERA  
EVAFHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRLMYQAGFARDQAREA  
YSGPDEDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRAL  
LTDPAHEVLVVTMGDRFGPHGAVGILLEKKPSTWHLKLLATSCRVVVSFGAGAT  
ILNWLTDQGARAGAHLVADFRRTDRNRMMEIAYRFAGFADSDCPCVSEVAGAS  
AAGVERLHLEPSARPAPTTLTLTAADIAPVTVSAAG.

For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkbl* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkbl* and *fkbl* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesize ethylmalonyl CoA or 2-hydroxymalonyl CoA. The invention provides methods and vectors for constructing recombinant *Streptomyces coelicolor* and *Streptomyces lividans* that are able to

synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

- In a preferred embodiment, the present invention provides recombinant
- 5 *Streptomyces* host cells, such as *S. coelicolor* and *S. lividans*, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that comprise one or more AT domains specific for ethylmalonyl CoA.
- 10 Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in which one or more AT domains is specific for ethylmalonyl CoA.

- In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For
- 15 example, deletion or inactivation of the *fkfG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkfG* gene product acts on 2-hydroxymalonyl and the resulting 2-
- 20 methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-506) are replaced by either hydrogen or methyl.

- This possibility of non-specific binding results from the construction of a hybrid
- 25 PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13,15-didesmethoxy-FK-520 and corresponding FK-506
- 30 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis, i.e., the *fkfH*, *fkfI*, *fkfJ*, and *fkfK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the

resulting host cells are thus preferred for production of polyketides that do not require the same.

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g.,  
5 U.S. Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-desmethoxy-FK-520; 13,15-didesmethoxy-FK-506; 13,15-  
10 didesmethoxy-FK-520; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-520; 13,15-didesmethoxy-18-hydroxy-FK-506; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,466; and 4,920,218, incorporated herein by reference.

15 Other compounds of the invention are shown in Figure 8, Parts A and B. In Figure 8, Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration and are applied to the dermis for treatment of conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in  
20 Figure 8, Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group, where R<sub>3</sub> and R<sub>4</sub> can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure  
25 8, Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15-methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

30 To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or  
35 triazole derivatives provides the C-32 tetrazole or teiazole derivative. As shown in the

lower scheme of Figure 8, Part B, reacting the starting compound with p-nitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, talc, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically,

parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

5 Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the  
10 present invention may be administered on an intermittent basis, i.e., at semi-weekly, weekly, semi-monthly, or monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral  
15 administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of,  
20 for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8% by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds  
25 of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular  
30 patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex, and diet of the subject; the time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

5

### Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-520

14

2

25

KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5           5'-CTAGTGGGCAGATCTGGCAGCT-3'  
           3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.

Next, a linker of the following sequence was ligated between the unique *Sph*I and *Afl*III sites of plasmid pKOS60-27-1 to introduce an *Nsi*I site at the 3' end of the module 8 cassette. The linker employed was:

10           5'-GGGATGCATGGC-3'  
           3'-GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (*Avr* II or *Nhe* I) and 3' end (*Xho* I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers *Spe*Bgl-fwd and either *Avr*-rev or *Nhe*-rev:

*Spe*Bgl-fwd   5'-CGACTCACTAGTGGGCAGATCTGG-3'  
           *Avr*-rev     5'-CACGCCTAGGCCGGTCTCGGGCCAC-3'  
20           *Nhe*-rev    5'-GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 µl reaction, 5 µl of 10x *Pfu* polymerase buffer (Stratagene), 5 µl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 µl DMSO, 2 µl of each primer (10 µM), 1 µl of template DNA (0.1 µg/µl), and 1 µl of cloned *Pfu* polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (*Bgl*II and *Avr*II or *Spe*I and *Nhe*I), and cloned into either pLitmus 28 or pLitmus38 (New England BioLabs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2, respectively.

30           Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers *Bsr*Xho-fwd and *Nsi*Afl-rev:

*Bsr*Xho-fwd   5'-GATGTACAGCTCGAGTCGGCACGCCCGGCCGCATC-3'  
           *Nsi*Afl-rev   5'-CGACTCACTTAAGCCATGCATCC-3'

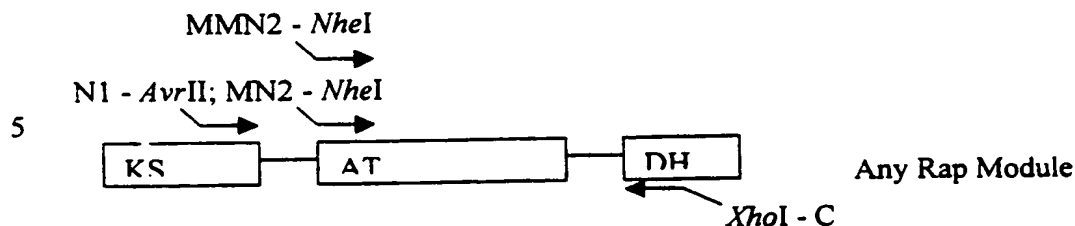
35           PCR conditions were as described above. The PCR fragment was cut with *Bsr*GI and *Afl*III, gel isolated, and ligated into pKOS60-37-4 cut with *Asp*718 and *Afl*III and



inserted into pKOS60-37-2 cut with *Bsr*GI and *Afl*III, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *Avr*II and *Xho*I or *Nhe*I and *Xho*I, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

- 5           Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *Avr*II or *Nhe*I site at the 5' end and an *Xho*I site at the 3' end. The PCR conditions were as given above and the primer sequences were as follows:

- 10   RATN1   5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3'  
          (3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA),  
      RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3'  
          (Rap AT shorter version 5'- sequence and specific for malonyl CoA),  
      RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3'  
15   (Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and  
      RATC   5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3'  
          (Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).



Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned PCR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The *AvrII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

```

20 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
   I W Q L A E A L L T L V R E S T
   GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
   A A V L G H V G G E D I P A T A A
   GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
25 F K D L G I D S L T A V Q L R N
   CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
   A L T E A T G V R L N A T A V F D
   TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250
   F P T P H V L A G K L G D E L T G
30 CACCCGCGCGCCCGTCTGTGCCCCGACCGCGGCCACGGCCGGTTCGACG 300
   T R A P V V P R T A A T A G A H
   ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350
   D E P L A I V G M A C R L P G G V
   GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
35 A S P E E L W H L V A S G T D A I
   CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450
   T E F P T D R G W D V D A I Y D
   CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
   P D P D A I G K T F V R H G G F L
40 ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550
   T G A T G F D A A F F G I S P R E
   GGCCCTCGCGATGGACCCGCAGCGGGTGCTCCTGGAGACGTCTGTGG 600
   A L A M D P Q Q R V L L E T S W
   AGGCGTTTCGAAAGCGCCGGCATACCCCGGACTCGACCCGCGGCAGCGAC 650
45 E A F E S A G I T P D S T R G S D
   ACCGGCGTGTTCGTGCGGCGCTTCTCCTACGGTTACGGCACCGGTGCGGA 700
   T G V F V G A F S Y G Y G T G A D
   CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750
   T D G F G A T G S Q T S V L S G
50 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800
   R L S Y F Y G L E G P A V T V D T
   GCGTGTTCGTGCTGCTGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850
   A C S S S L V A L H Q A G Q S L R

```

CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900  
S G E C S L A L V G G V T V M A  
CTCCCGGCGGTTCTCGTGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950  
S P G G F V E F S R Q R G L A P D  
5 GGCCGGCGGAAGGCGTTCGGCGCGGGTGCGGACGGCACGAGCTTCGCCGA 1000  
G R A K A F G A G A D G T S F A E  
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
G A G V L I V E R L S D A E R N  
GTCACACCGTCTGGCGGTTCGTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100  
10 G H T V L A V V R G S A V N Q D G  
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTTCGAGGAGCGGGTGAT 1150  
A S N G L S A P N G P S Q E R V I  
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGGACGTGGACGCCG 1200  
R Q A L A N A G L T P A D V D A  
15 TCGAGGCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G T G T R L G D P I E A Q  
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCTGCTGCTGGG 1300  
A V L A T Y G Q E R A T P L L L G  
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCGCCG 1350  
20 S L K S N I G H A Q A A S G V A  
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
G I I K M V Q A L R H G E L P P T  
CTGCACGCCGACGAGCCGTGCGCGCACGTTCGACTGGACGGCCGGCGCCGT 1450  
L H A D E P S P H V D W T A G A V  
25 CGAACTGCTGACGTGCGCCCGGCCGTGGCCCGAGACCGACCGGCCTAGGC 1500  
E L L T S A R P W P E T D R P R  
GGGCAGGCGTGTGCTCCTTCGGGATCAGTGGCACCAACGCCACGTCATC 1550  
R A G V S S F G I S G T N A H V I  
CTGGAAAGCGCACCCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG 1600  
30 L E S A P P T Q P A D N A V I E R  
GGCACCGGAGTGGGTGCCGTTGGTGATTTCCGGCCAGGACCCAGTCGGCTT 1650  
A P E W V P L V I S A R T Q S A  
TGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTGGCGGCGTCGCCCGGG 1700  
L T E H E G R L R A Y L A A S P G  
35 GTGGATATGCGGGCTGTGGCATCGACGCTGGCGATGACACGGTCGGTGTT 1750  
V D M R A V A S T L A M T R S V F  
CGAGCACCGTGCCGTGTGTTGGGAGATGACACCGTCACCGGCACCGCTG 1800  
E H R A V L L G D D T V T G T A  
TGTCTGACCCTCGGGCGGTGTTCTGCTCTTCCCGGGACAGGGGTTCGACGCT 1850  
40 V S D P R A V F V F P G Q G S Q R  
GCTGGCATGGGTGAGGAACCTGGCCGCCCGGTTCCCGTCTTCGCGCGGAT 1900  
A G M G E E L A A A F P V F A R I  
CCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCGATCTGGAGGTGAACG 1950  
H Q Q V W D L L D V P D L E V N  
45 AGACCGGTTACGCCAGCCGGCCCTGTTTCGAATGCAGGTGGCTCTGTTC 2000  
E T G Y A Q P A L F A M Q V A L F  
GGGCTGCTGGAATCGTGGGTGTACGACCGGACGCGGTGATCGGCCATTG 2050  
G L L E S W G V R P D A V I G H S  
GGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGTGTGGTTCGTTGGAGG 2100  
50 V G E L A A A Y V S G V W S L E  
ATGCCTGCACTTTGGTGTGCGCGCGGGCTCGTCTGATGCAGGCTCTGCCC 2150  
D A C T L V S A R A R L M Q A L P  
GCGGGTGGGGTGATGGTTCGTGTCCCGGTCTCGGAGGATGAGGCCCGGGC 2200  
A G G V M V A V P V S E D E A R A  
55 CGTGCTGGGTGAGGGTGTGGAGATCGCCCGGTCAACGGCCCGTCTGTCGG 2250  
V L G E G V E I A A V N G P S S  
TGTTCTCTCCGGTGATGAGGCCGCCGTGCTGCAGGCCGCGGAGGGGCTG 2300  
V V L S G D E A A V L Q A A E G L  
GGGAAGTGGACGCGGCTGGCGACCGACCGGTTCCATTCCGCCCGTAT 2350  
60 G K W T R L A T S H A F H S A R M  
GGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCGAAGGCCTGACCTACC 2400  
E P M L E E F R A V A E G L T Y  
GGACGCCGAGGTCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAG 2450  
R T P Q V S M A V G D Q V T T A E

TACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGC 2500  
Y W V R Q V R D T V R F G E Q V A  
CTCGTACGAGGACGCCGTGTTTCGTCGAGCTGGGTGCCGACCGGTCACTGG 2550  
S Y E D A V F V E L G A D R S L  
5 CCGGCCTGGTCGACGGTGTGCGGATGCTGCACGGCGACCACGAAATCCAG 2600  
A R L V D G V A M L H G D H E I Q  
GCCGCGATCGGCGCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCTGA 2650  
A A I G A L A H L Y V N G V T V D  
10 CTGGCCCGCGCTCCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTC 2700  
W P A L L G D A P A T R V L D L  
CGACATACGCCTTCCAGCACCAGCGCTACTGGCTCGAGTCGGCACGCCCG 2750  
P T Y A F Q H Q R Y W L E S A R P  
GCCGCGATCCGACGCGGGCCACCCCGTGGTGGCTCCGGTATCGCCCTCGC 2800  
A A S D A G H P V L G S G I A L A  
15 CCGGTGCGCCGGGCGGGTGTTCACGGGTTCGTCGCGACCGGTGCGGACC 2850  
G S P G R V F T G S V P T G A D  
GCGCGGTGTTCTGTCGCGAGCTGGCGCTGGCCGCGCGGACGCGGTGCGAC 2900  
R A V F V A E L A L A A A D A V D  
TGCGCCACGGTCGAGCGGCTCGACATCGCCTCCGTGCCCGGCGCGCGGG 2950  
20 C A T V E R L D I A S V P G R P G  
CCATGGCCGGGACGACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGACG 3000  
H G R T T V Q T W V D E P A D D  
GCCGGCGCGGGTTCACCGTGCACACCCGACCGGCGACGCCCCGTGGACG 3050  
G R R F T V H T R T G D A P W T  
25 CTGCACGCGGAGGGGGTGTGCGCCCCATGGCACGGCCCTGCCCGATGC 3100  
L H A E G V L R P H G T A L P D A  
GGCCGACGCGGAGTGGCCCCCACCAGGGCGCGGTGCCCGCGGACGGGTGC 3150  
A D A E W P P P G A V P A D G L  
CGGGTGTGTGGCGCCGGGGGACCAGGTCTTCGCGGAGGCCGAGGTGGAC 3200  
30 P G V W R R G D Q V F A E A E V D  
GGACCGGACGGTTCGTTGTCACCCCGACCTGCTCGACGCGGTCTTCTC 3250  
G P D G F V V H P D L L D A V F S  
CGCGGTGCGCGACGGAAGCCGCCAGCCGGCCGATGGCGCGACCTGACGG 3300  
A V G D G S R Q P A G W R D L T  
35 TGCACGCGTCGGACGCCACCGTACTGCGCGCCTGCCTCACC CGGCGCACC 3350  
V H A S D A T V L R A C L T R R T  
GACGGAGCCATGGGATTGCGCGCCTTCGACGCGCGCCGCGCTGCCGGTACT 3400  
D G A M G A A F D G A G L P V L  
CACCGCGGAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCG 3450  
40 T A E A V T L R E V A S P S G S  
AGGAGTCGGACGGCCTGCACCGTTGGAGTGGCTCGCGGTGCGCGGAGGCG 3500  
E E S D G L H R L E W L A V A E A  
GTCTACGACGGTGACCTGCCCCGAGGGACATGTCCTGATCACCGCCGCCCA 3550  
V Y D G D L P E G H V L I T A A H  
45 CCCCAGACGCCCGAGGACATACCCACCCGCGCCACACCCGCGCCACCC 3600  
P D D P E D I P T R A H T R A T  
GCGTCTGACCGCCCTGCAACACCACCTCACCACCACCGACACACCCCTC 3650  
R V L T A L Q H H L T T T D H T L  
ATCGTCCACACCACCGACCCCGCGCGGCCACCGTCACCGGCCTCAC 3700  
50 I V H T T T D P A G A T V T G L T  
CCGACCGCCAGAACGAACACCCCGACCGCATCCGCCTCATCGAAACCG 3750  
R T A Q N E H P H R I R L I E T  
ACCACCCCGACACCCCGCTCCCGCTGGCCCAACTCGCCACCCTCGACCAC 3800  
D H P H T P L P L A Q L A T L D H  
55 CCCCACCTCCGCTCACCCACCACCCCTCCACCACCCCGACCTCACCC 3850  
P H L R L T H H T L H H P H L T P  
CCTCCACACCACCCCGACCCACCCACCCCGCTCAACCCCGAACACG 3900  
L H T T T P P T T T P L N P E H  
CCATCATCATCACGGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGCCCGC 3950  
60 A I I I T G G S G T L A G I L A R  
CACCTGAACACCCCGACACCTACCTCCTCCCGCACCCCGACCCCGCA 4000  
H L N H P H T Y L L S R T P P P D  
CGCCACCCCGGCGACCCACCTCCCGTGGGACGTGGGCGACCCCGACCAAC 4050  
A T P G T H L P C D V G D P H Q

TCGCCACCACCTCACCACATCCCCAACCCTCACCGCCATCTTCCAC 4100  
 L A T T L T H I P Q P L T A I F H  
 ACCGCCGCCACCTCGACGACGGCATCTCCACGCCCTCACCCCGACCG 4150  
 T A A T L D D G I L H A L T P D R  
 5 CCTCACCACCGTCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACC 4200  
 L T T V L H P K A N A A W H L H  
 ACCTCACCACCAACCCCTCACCACCTTCGTCTCTACTCCAGCGCC 4250  
 H L T Q N Q P L T H F V L Y S S A  
 GCCGCCGTCTCGGCAGCCCCGGACAAGGAACTACGCCGCCGCCAACGC 4300  
 10 A A V L G S P G Q G N Y A A A N A  
 CTTCTCGACGCCCTCGCCACCCACGCCACCCCTCGGCCAACCCGCCA 4350  
 F L D A L A T H R H T L G Q P A  
 CCTCCATCGCCTGGGGCATGTGGCACACCACAGCACCTCACCAGGACAA 4400  
 T S I A W G M W H T T S T L T G Q  
 15 CTCGACGACGCCGACGGGACCGCATCCGCCGCGGGGTTTCTCCCGAT 4450  
 L D D A D R D R I R R G G F L P I  
 CACGGACGACGAGGGCATGGGGATGCAT  
 T D D E G

- 20 The *AvrII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 25 Q L A E A L L T L V R E S T  
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100  
 A A V L G H V G G E D I P A T A A  
 GTTCAAGGACCTCGGCATCGACTCGCTCACC CGGTCCAGCTGCGCAACG 150  
 F K D L G I D S L T A V Q L R N  
 30 CCCTCACCAGAGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAAGTACCGG 250  
 F P T P H V L A G K L G D E L T G  
 CACCCGCGCGCTCGTCTGCCCCGACCGCGGCCACGGCCCGGTGCGCACG 300  
 35 T R A P V V P R T A A T A G A H  
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGCGTCCCGCGGGGTC 350  
 D E P L A I V G M A C R L P G G V  
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400  
 A S P E E L W H L V A S G T D A I  
 40 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450  
 T E F P T D R G W D V D A I Y D  
 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCTC 500  
 P D P D A I G K T F V R H G G F L  
 ACCGGCGGACAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550  
 45 T G A T G F D A A F F G I S P R E  
 GGCCCTCGCGATGGACCCGACGAGCGGGTGTCTCTGGAGACGTCTGTGG 600  
 A L A M D P Q Q R V L L E T S W  
 AGGCGTTGAAAGCGCCGGCATCACC CGGACTCGACCCGCGGCAGCGAC 650  
 E A F E S A G I T P D S T R G S D  
 50 ACCGGCGTGTTCGTGCGCGCCTTCTCTACGTTACGGCACCGGTGCGGA 700  
 T G V F V G A F S Y G Y G T G A D  
 CACCGACGGCTTCGGCGGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750  
 T D G F G A T G S Q T S V L S G  
 GGCTGTGCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800  
 55 R L S Y F Y G L E G P A V T V D T  
 GCGTGTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGCAGTCTGCTGCG 850  
 A C S S S L V A L H Q A G Q S L R  
 CTCCGGCGAATGCTCGCTCGCCCTGGTCCGGCGGCTCACGGTGATGGCGT 900  
 S G E C S L A L V G G V T V M A  
 60 CTCCCGCGGCTTCGTGGAGTTCTCCCGGACGCGGGCTCGCGCCGGAC 950

S P G G F V E F S R Q R G L A P D  
GGCCGGGCGAAGGCGTTCGGCGCGGGTGGCGACGGCAGAGCTTCGCCGA 1000  
G R A K A F G A G A D G T S F A E  
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
5 G A G V L I V E R L S D A E R N  
GTCACACCGTCTTGGCGGTCTGCTGGTTCGGCGGTCAACCAGGATGGT 1100  
G H T V L A V V R G S A V N Q D G  
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTGCGAGGAGCGGGTGAT 1150  
A S N G L S A P N G P S Q E R V I  
10 CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCGGCGGACGTGGACGCCG 1200  
R Q A L A N A G L T P A D V D A  
TCGAGGCCCCACGGCACCGGCACCGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G C T G T R L G D P I E A Q  
GCGGTACTGGCCACCTACGGACAGCGCGCCACCCCTGCTGCTGGG 1300  
15 A V L A T Y G Q E R A T P L L L G  
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCGCCG 1350  
S L K S N I G H A Q A A S G V A  
GCATCATCAAGATGGTGCAGGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
G I I K M V Q A L R H G E L P P T  
20 CTGCACGCCGACGAGCCGTCGCCGACGTCGACTGGACGGCCGGCGCCGT 1450  
L H A D E P S P H V D W T A G A V  
CGAACTGCTGACGTCGGCCCCGGCCGTGGCCCCGAGACCGACCGGCCTAGGC 1500  
E L L T S A R P W P E T D R P R  
GGGCGGGCGTGTGCTCCTTCGGAGTCAGCGGCACCAACGCCACGTCATC 1550  
25 R A G V S S F G V S G T N A H V I  
CTGGAGAGCGCACCCCCGCTCAGCCCCGGGAGGAGGCGCAGCCTGTTGA 1600  
L E S A P P A Q P A E E A Q P V E  
GACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGGTGATATCGGCCAAGA 1650  
T P V V A S D V L P L V I S A K  
30 CCCAGCCCCGCTGACCGAACACGAAGACCGGCTGCGCGCTACCTGGCG 1700  
T Q P A L T E H E D R L R A Y L A  
GCGTCGCCCCGGGGCGGATATACGGGCTGTGGCATCGACGCTGGCGGTGAC 1750  
A S P G A D I R A V A S T L A V T  
ACGGTCGGTGTTCGAGCACCGCGCCGACTCCTTGGAGATGACACCGTCA 1800  
35 R S V F E H R A V L L G D D T V  
CCGGCACCGCGGTGACCGACCCAGGATCGTGTGTTGCTTTCCCGGGCAG 1850  
T G T A V T D P R I V F V F P G Q  
GGGTGGCAGTGCTGGGGATGGGCAGTGCCTGCGCGATTCTGTCGGTGGT 1900  
G W Q W L G M G S A L R D S S V V  
40 GTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCGTTGCGCGAGTTCGTGG 1950  
F A E R M A E C A A A L R E F V  
ACTGGGATCTGTTACGGTTCTGGATGATCCGGCGGTGGTGGACCGGGTT 2000  
D W D L F T V L D D P A V V D R V  
GATGTGGTCCAGCCCGCTTCTGCGGATGATGGTTTCCCTGGCCGCGGT 2050  
45 D V V Q P A S W A M M V S L A A V  
GTGGCAGGCGGCGGTGTGCGGCCGGATGCGGTGATCGGCCATTTCGACAG 2100  
W Q A G V R P D A V I G H S Q  
GTGAGATCGCCGAGCTTGTGTGGCGGGTGGGTGTCACTACGCGATGCC 2150  
G E I A A A C V A G A V S L R D A  
50 GCCCGGATCGTGACCTTGGCGAGCCAGGCGATCGCCCGGGGCTGGCGGG 2200  
A R I V T L R S Q A I A R G L A G  
CCGGGGCGCGATGGCATCCGTCGCCCTGCCGCGCAGGATGTGAGCTGG 2250  
R G A M A S V A L P A Q D V E L  
TCGACGGGGCTGGATCGCCGCCCAACGGGCGCCCTCCACCGTGATC 2300  
55 V D G A W I A A H N G P A S T V I  
GCGGGCACCCCGGAAGCGGTGACCATGTCCTACCGCTCATGAGGCACA 2350  
A G T P E A V D H V L T A H E A Q  
AGGGGTGCGGGTGGCGCGGATCACCGTCGACTATGCCTCGCACACCCCGC 2400  
G V R V R R I T V D Y A S H T P  
60 ACGTCGAGCTGATCCGCGACGAATACTCGACATCACTAGCGACAGCAGC 2450  
H V E L I R D E L L D I T S D S S  
TCGACAGACCCCGCTCGTGCCGTGGCTGTGACCGTGGACGGCACCTGGGT 2500  
S Q T P L V P W L S T V D G T W V  
CGACAGCCCGCTGGACGGGGAGTACTGGTACCGGAACCTGCGTGAACCGG 2550

D S P L D G E Y W Y R N L R E P  
 TCGGTTTCCACCCCGCCGTCAGCCAGTTGACAGGCCAGGGCGACACCGTG 2600  
 V' G F H P A V S Q L Q A Q G D T V  
 TTCGTCGAGGTCAGCGCCAGCCCGGTGTTGTTGCAGGCGATGGACGACGA 2650  
 5 F V E V S A S P V L L Q A M D D D  
 TGTCGTCACGGTTGCCACGCTGCGTCTGACGACGGCGACGCCACCCGGA 2700  
 V V T V A T L R R D D G D A T R  
 TGCTCACCGCCCTGGCACAGGCCTATGTCCACGGCGTCACCGTCGACTGG 2750  
 M L T A L A Q A Y V H G V T V D W  
 10 CCCGCCATCCTCGGCACCACCACAACCCGGTACTGGACCTTCCGACCTA 2800  
 P A I L G T T T T R V L D L P T Y  
 CGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGGCACGCCCCGGCCGAT 2850  
 A F Q H Q R Y W L E S A R P A A  
 CCGACGCGGGGCCACCCCGTGCTGGGCTCCGGTATCGCCCTCGCCGGGTGCG 2900  
 15 S D A G H P V L G S G I A L A G S  
 CCGGGCCGGGTGTTACGGGTCCGTGCCGACCGGTGCGGACCGCGCGGT 2950  
 P G R V F T G S V P T G A D R A V  
 GTTCGTCGCCGAGCTGGCGCTGGCCGCCGCGACGCGGTGCGACTGCGCCA 3000  
 F V A E L A L A A A D A V D C A  
 20 CGGTGCGAGCGGCTCGACATCGCCTCCGTGCCCGGCCGGCCGGGCCATGGC 3050  
 T V E K L D I A S V P G R P G H G  
 CGGACGACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGACGGCCGGCG 3100  
 R T T V Q T W V D E P A D D G R R  
 CCGGTTACCGTGACACCCGACCGGCGACGCCCCGTGGACGCTGCACG 3150  
 25 R F T V H T R T G D A P W T L H  
 CCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGCGGCCGAC 3200  
 A E G V L R P H G T A L P D A A D  
 GCGAGTGCCCCCACC GGCGCGGTGCCCGCGGACGGGCTGCCGGGTGT 3250  
 A E W P P P G A V P A D G L P G V  
 30 GTGGCGCCGGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGACGGACCGG 3300  
 W R R G D Q V F A E A E V D G P  
 ACGGTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTC 3350  
 D G F V V H P D L L D A V F S A V  
 GGCGACGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGACGC 3400  
 35 G D G S R A P A G W R D L T V H A  
 GTCGGACGCCACCGTACTGCGCGCCTGCCTACCCGGCGCACCGACGGAG 3450  
 S D A T V L R A C L T R R T D G  
 CCATGGGATTGCGCGCCTTCGACGGCGCCGGCCTGCCGGTACTCACCGCG 3500  
 A M G F A A F D G A G L P V L T A  
 40 GAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTC 3550  
 E A V T L R E V A S P S G S E E S  
 GGACGGCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACG 3600  
 D G L H R L E W L A V A E A V Y  
 ACGGTGACCTGCCCGAGGGACATGTCTGATCACCGCCGCCCCACCCCGAC 3650  
 45 D G D L P E G H V L I T A A H P D  
 GACCCCGAGGACATACCCACCCGCGCCACACCCGCGCCACCCGCGTCCT 3700  
 D P E D I P T R A H T R A T R V L  
 GACCGCCCTGCAACACCACCTCACCACCACCGACACACCTCATCGTCC 3750  
 T A L Q H H L T T T D H T L I V  
 50 ACACCACCACCGACCCCGCCGGCGCCACCGTCACCGGCCTCACCCGCACC 3800  
 H T T T D P A G A T V T G L T R T  
 GCCCAGAACGAACACCCCCACCGCATCCGCCTCATCGAAACCGACACCC 3850  
 A Q N E H P H R I R L I E T D H P  
 CCACACCCCCCTCCCCCTGGCCCAACTCGCCACCCCTCGACCACCCCCACC 3900  
 55 H T P L A Q L A T L D H P H  
 TCCGCCTCACCCACCACACCCCTCCACCACCCCCACCTCACCCCCCTCCAC 3950  
 L R L T H H T L H H P H L T P L H  
 ACCACCACCCACCCACCACCCACCCCTCAACCCGAACACGCCATCAT 4000  
 T T T P P T T T P L N P E H A I I  
 60 CATCACCGGCGGCTCCGGCACCCCTCGCCGGCATCTCGCCCGCCACCTGA 4050  
 I T G G S G T L A G I L A R H L  
 ACCACCCACACCTACCTCTCCCGCACCCACCCCCGACGCCACC 4100  
 N H P H T T Y L L S R T P P P D A T  
 CCCGACCCACCTCCCCTGCGACGTGGCGACCCCCACCAACTCGCCAC 4150

P G T H L P C D V G D P H Q L A T  
 CACCCTCACCCACATCCCCCAACCCCTCACCGCCATCTTCCACACCGCCG 4200  
 T L T H I P Q P L T A I F H T A  
 CCACCCTCGACGACGGCATCCTCCACGCCCTCACCCCCGACCGCCTCACC 4250  
 5 A T L D D D G I L H A L T P D R L T  
 ACCGTCCTCCACCCCAAGCCAACGCCGCTGGCACCTGCACCACCTCAC 4300  
 T V L H P K A N A A W H L H H L T  
 CCAAACCAACCCCTCACCCTTCTGTCCTCTACTCCAGCGCCGCCGCCG 4350  
 Q N Q P L T H F V L Y S S A A A  
 10 TCCTCGGCAGCCCCGGACAAGGAACTACGCCGCCGCCAACGCCTTCCTC 4400  
 V L G S P G Q G N Y A A A N A F L  
 GAGCCCTCGCCACCCACCGCCACCCCTCGGCCAACCCGCCACCTCCAT 4450  
 D A L A T H R H T L G Q P A T S I  
 CGCCTGGGGCATGTGGCACACCACCAGCACCTCACCAGCAACTCGACG 4500  
 15 A W G M W H T T S T L T G Q L D  
 ACGCCGACCGGGACCGCATCCGCCCGGGCGGTTTCTCCCGATCACGGAC 4550  
 D A D R D R I R R G G F L P I T D  
 GACGAGGGCATGGGGATGCAT  
 D E G  
 20

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 (specific for malonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

25 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 Q L A E A L L T L V R E S T  
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100  
 A A V L G H V G G E D I P A T A A  
 GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150  
 30 F K D L G I D S L T A V Q L R N  
 CCCTCACCGAGGCGACCGGTGTGCGGTGAACGCCACGGCGGTCTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGGCAGCAACTGACCGG 250  
 F P T P H V L A G K L G D E L T G  
 35 CACCCGCGCGCCCGTCTGTGCCCCGACCGCGGCCACGGCCGCTGCGCACG 300  
 T R A P V V P R T A A T A G A H  
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350  
 D E P L A I V G M A C R L P G G V  
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCAGCGCCAT 400  
 40 A S P E E L W H L V A S G T D A I  
 CACGGAGTTCCCGACGGACCGCGGTGGGACGTGACGCGATCTACGACC 450  
 T E F P T D R G W D V D A I Y D  
 CGGACCCCGACCGATCGGCAAGACCTTCGTCCGGCAGGTTGGCTTCCTC 500  
 P D P D A I G K T F V R H G G F L  
 45 ACCGGCGCGACAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550  
 T G A T G F D A A F F G I S P R E  
 GGCCCTCGCGATGGACCCGACGAGCGGGTGCTCCTGGAGACGTGCTGGG 600  
 A L A M D P Q Q R V L L E T S W  
 AGGCGTTCAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650  
 50 E A F E S A G I T P D S T R G S D  
 ACCGGCGTGTTCGTGCGCGCCTTCTCTACGGTTACGGCACCGGTGCGGA 700  
 T G V F V G A F S Y G Y G T G A D  
 CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750  
 T D G F G A T G S Q T S V L S G  
 55 GGCTGTGCTACTTCTACGGTCTGGAGGTCGGCGGTCACGGTCGACACG 800  
 R L S Y F Y G L E G P A V T V D T  
 GCGTGTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850  
 C S S L V A L H Q A G Q S L R  
 CTCCGGCGAATGCTCGCTCGCCCTGGTTCGGCGGCGTCACGGTGATGGCGT 900  
 60 S G E C S L A L V G G V T V M A



CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950  
S P G G F V E F S R Q R G L A P D  
GGCCGGGCGAAGGCGTTCCGGCGGGTGCGGACGGCAGAGCTTCGCCGA 1000  
G R A K A F G A G A D G T S F A E  
5 GGGTGCCGGTGTGCTGATCGTTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
G A G V L I V E R L S D A E R N  
GTCACACCGTCTCGGCGGTGTCGCGGTCAACCAGGATGGT 1100  
G H T V L A V V R G S A V N Q D G  
GCCTCCAACGCGCTGTGCGGCGCGAACGGGCGCTCGCAGGAGCGGGTGAT 1150  
10 A S N G L S A P N G P S Q E R V I  
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGGACGTGGACGCCG 1200  
R Q A L A N A G L T P A D V D A  
TCGAGGCCCCACGGCACCGGCACCGAGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G T G T R L G D P I E A Q  
15 GCGGTACTGGCCACTACGGACAGGAGCGCGCCACCCCTGCTGCTGGG 1300  
A V L A T Y G Q E R A T P L L L G  
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCCGCCG 1350  
S L K S N I G H A Q A A S G V A  
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
20 G I I K M V Q A L R H G E L P P T  
CTGCACGCCGACGAGCCGTGCGCCGACGTGACTGGACGGCCGGCGCCGT 1450  
L H A D E P S P H V D W T A G A V  
CGAACTGCTGACGTGCGCCCGGCCGTGGCCCGAGACCGACCGGCCACGGC 1500  
E L L T S A R P W P E T D R P R  
25 GTGCCGCCGTCTCCTCGTTCCGGGTGAGCGGCACCAACGCCACGTCATC 1550  
R A A V S S F G V S G T N A H V I  
CTGGAGGCCGACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGA 1600  
L E A C P V T E T P A A S P S G D  
CCTTCCCTGCTGGTGTGCGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650  
30 L P L L V S A R S P E A L D E Q  
TCCGCCGACTGCGCGCCTACCTGGACACCACCCCGGACGTGACCGGGTG 1700  
I R R L R A Y L D T T P D V D R V  
GCCGTGGCACAGACGCTGGCCCGGCGCACACTTCGCCACCGCGCCGT 1750  
A V A Q T L A R R T H F A H R A V  
35 GCTGCTCGGTGACACCGTCATCACACACCCCGCGGACCGGCCGACG 1800  
L L G D T V I T T P P A D R P D  
AACTCGTCTTCGTCTACTCCGGCCAGGGCACCCAGCATCCCGCGATGGGC 1850  
E L V F V Y S G Q G T Q H P A M G  
GAGCAGCTAGCCGCCGCGTTCCTCCCGTCTTCGCGCGGATCCATCAGCAGGT 1900  
40 E Q L A T A A F P V F A R I H Q Q V  
GTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACGAGACCGGTTACG 1950  
W D L L D V P D L E V N E T G Y  
CCCAGCCGGCCCTGTTCCGAATGCAGGTGGCTCTGTTCCGGGCTGCTGGAA 2000  
A Q P A L F A M Q V A L F G L L E  
45 TCGTGGGGTGTACGACCGGACGCGGTGATCGGCCATTCCGTGGGTGAGCT 2050  
S W G V R P D A V I G H S V G E L  
TGCGGCTGCGTATGTGTCCGGGGTGTGGTTCGTTGGAGGATGCCTGCACTT 2100  
A A A Y V S G V W S L E D A C T  
TGGTGTGCGGCGGGGCTCGTCTGATGCAGGCTCTGCCCGCGGGTGGGGTG 2150  
50 L V S A R A R L M Q A L P A G G V  
ATGGTTCGCTGTCCCGGTCTCGGAGGATGAGGCCCGGGCCGTGCTGGGTGA 2200  
M V A V P V S E D E A R A V L G E  
GGGTGTGAGATCGCCGCGGTCAACGGCCCGTCTGTCGGTGGTTCTCTCCG 2250  
G V E I A A V N G P S S V V L S  
55 GTGATGAGGCCGCGGTGCTGTCAGGCGCGGAGGGGCTGGGGAAGTGGACG 2300  
G D E A A V L Q A A E G L G K W T  
CGGCTGGCGACCGACGCGTTCATTCCGCCCGTATGGAACCCATGCT 2350  
R L A T S H A F H S A R M E P M L  
GGAGGAGTTCGGGCGGTGCGCCGAAGGCCTGACCTACCGGACGCCGACG 2400  
60 E F R A V A E G L T Y R T P Q  
TCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450  
V S M A V G D Q V T T A E Y W V R  
CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500  
Q V R D T V R F G E Q V A S Y E D

CGCCGTGTTTCGTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCCTGGTTCG 2550  
A V F V E L G A D R S L A R L V  
ACGGTGTGCGATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGC 2600  
D G V A M L H G D H E I Q A A I G  
5 GCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCGACTGGCCCGCGCT 2650  
A L A H L Y V N G V T V D W P A L  
CCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCT 2700  
L G D A P A T R V L D L P T Y A  
TCCAGCACCAGCGCTACTGGCTCGAGTCGGCACGCCCCGCCGATCCGAC 2750  
10 F Q H Q R Y W L E S A R P A A S D  
GCGGGCCACCCCGTGTGCTGGGCTCCGGTATCGCCCTCGCCGGGTGCGCGGG 2800  
A G H P V L G S G I A L A G S P G  
CCGGGTGTTACGGGTTCGGTGCCGACCGGTGCGGACCGCGCGGTGTTTCG 2850  
R V F T G S V P T G A D R A V F  
15 TCGCCGAGCTGGCGCTGGCCGCCGCGGACGCGGTGCGACTGCGCCACGGTC 2900  
V A E L A L A A A D A V D C A T V  
GAGCGGCTCGACATCGCCTCCGTGCCCCGCCGCGGGCCATGGCCGGAC 2950  
E R L D I A S V P G R P G H G R T  
GACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGCGCCGCGCGCGGT 3000  
20 T V Q T W V D E P A D D G R R R  
TCACCGTGCACACCCGACCGCGACGCCCCGTGGACGCTGCACGCCGAG 3050  
F T V H T R T G D A P W T L H A E  
GGGGTGCTCGCCCCCATGGCACGGCCCTGCCGATGCGGGCCGACGCCGA 3100  
G V L R P H G T A L P D A A D A E  
25 GTGGCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGCCGGGTGTGTGGC 3150  
W P P P G A V P A D G L P G V W  
GCCGGGGGACAGGTCTTCGCGGAGGCCGAGGTGGACGGACCGGACGGT 3200  
R R G D Q V F A E A E V D G P D G  
TTCGTGGTGCAACCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGCGA 3250  
30 F V V H P D L L D A V F S A V G D  
CGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGCGTCGG 3300  
G S R Q P A G W R D L T V H A S  
ACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCGACGGAGCCATG 3350  
D A T V L R A C L T R R T D G A M  
35 GGATTGCGCGCCTTCGACGGCGCCGGCCTGCCGGTACTCACCGCGGAGGC 3400  
G F A A F D G A G L P V L T A E A  
GGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTCGGACG 3450  
V T L R E V A S P S G S E E S D  
GCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACGACGGT 3500  
40 G L H R L E A V A E A V Y D G  
GACCTGCCCCGAGGACATGTCTGATCACCGCCGCCACCCCGACGACCC 3550  
D L P E G H V L I T A A H P D D P  
CGAGGACATACCCACCCGCGCCACACCCGCGCCACCCGCGTCTTGACCG 3600  
E D I P T R A H T R A T R V L T  
45 CCCTGCAACACCACCTCACCACCACCGACCACCCCTCATCGTCCACACC 3650  
A L Q H H L T T T D H T L I V H T  
ACCACGACCCCGCGGCCACCGTACCGGCCTCACCCGCACCGCCCA 3700  
T T D P A G A T V T G L T R T A Q  
GAACGAACACCCCCACCGCATCCGCTCATCGAAACCGACACCCCCACA 3750  
50 N E H P H R I R L I E T D H P H  
CCCCCTCCCCCTGGCCCAACTCGCCACCCTCGACCACCCCCACCTCCGC 3800  
T P L P L A Q L A T L D H P H L R  
CTCACCCACCAACCCCTCCACCACCCCACTCACCCCTCCACACCAC 3850  
L T H H T L H P H L T P L H T T  
55 CACCCACCCACCAACCCCTCAACCCGAACACGCCATCATCATCA 3900  
T P P T T T P L N P E H A I I I  
CCGGCGGTCCGGCACCTCGCCGGCATCCTCGCCCGCCACCTGAACCAC 3950  
T G G S G T L A G I L A R H L N H  
CCCCACACCTACCTCCTCTCCCGCACCCACCCCGGACGCCACCCCGG 4000  
60 P H T Y L L S R T P P P D A T P G  
CACCCACCTCCCCTGCGACGTGGGCGACCCCACTCGCCACCAACC 4050  
T H L P C D V G D P H Q L A T T  
TCACCCACATCCCCCAACCCCTCACCGCATCTTCCACACCGCGCCACC 4100  
L T H I P Q P L T A I F H T A A T

CTCGACGACGGCATCCTCCACGCCCTCACCCCGACCGCTCACCACCGT 4150  
 L D D G I L H A L T P D R L T T V  
 CCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACCACCTCACCCAAA 4200  
 L H P K A N A A W H L H H L T Q  
 5 ACCAACCCCTCACCACTTCTGCTCTACTCCAGCGCGCGCGCTCCTC 4250  
 N Q P L T H F V L Y S S A A A V L  
 GGCAGCCCGGACAAGGAACTACGCCGCGCCAAACGCCTTCCTCGACGC 4300  
 G S P G Q G N Y A A A N A F L D A  
 CCTCGCCACCCACCGCCACACCCTCGGCCAACCGCCACCTCCATCGCCT 4350  
 10 L A T H R H T L G Q P A T S I A  
 GGGGCATGTGGCACACCACCAGCACCTCACCGGACAACCTCGACGACGCC 4400  
 W G N W H T T S T L T G Q L D D A  
 GACCGGGACCGCATCCGCCGCGCGGTTTCTCCCGATCACGGACGACGA 4450  
 D R D R I R R G G F L P I T D D E  
 15 GGGCATGGGGATGCAT  
 G

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS  
 with the endogenous AT domain replaced by the AT domain of module 13 (specific for  
 20 methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the  
 amino acid sequence shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 Q L A E A L L T L V R E S T  
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100  
 25 A A V L G H V G G E D I P A T A A  
 GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150  
 F K D L G I D S L T A V Q L R N  
 CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 30 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACCTGACCGG 250  
 F P T P H V L A G K L G D E L T G  
 CACCGCGCGCGCGTCTGCCCCGGACCGCGGCCACGGCCGGTTCGCGACG 300  
 T R A P V V P R T A A T A G A H  
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGGGGGTC 350  
 35 D E P L A I V G M A C R L P G G V  
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400  
 A S P E E L W H L V A S G T D A I  
 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450  
 T E F P T D R G W D V D A I Y D  
 40 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500  
 P D P D A I G K T F V R H G G F L  
 ACCGGCGCGACAGGCTTCGACGCGCGTTCCTCGGCATCAGCCCGCGCGA 550  
 T G A T G F D A A F F G I S P R E  
 GGCCCTCGCGATGGACCCGCGAGCAGCGGGTGTCTCTGGAGACGTCTGTTGG 600  
 45 A L A M D P Q Q R V L L E T S W  
 AGGCGTTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCGAGCGAC 650  
 E A F E S A G I T P D S T R G S D  
 ACCGGCGTGTTCGTGCGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700  
 T G V F V G A F S Y G Y G T G A D  
 50 CACCGACGGCTTCGGCGCGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750  
 T D G F G A T G S Q T S V L S G  
 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTCGACACG 800  
 R L S Y F Y G L E G P A V T V D T  
 GCGTGTTCGTCTGCTGGTGGCGCTGCACCAGGCCGGGCGAGTCGCTGCG 850  
 55 A C S S S L V A L H Q A G Q S L R  
 CTCCGGCGAATGCTCGCTCGCCCTGGTCCGGCGGCGTCACGGTGATGGCGT 900  
 S G E C S L A L V G G V T V M A  
 CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCGAGCGCGGCTCGCGCCGGAC 950  
 S P G G F V E F S R Q R G L A P D  
 60 GGCCGGTCGAAGGCGTTTCGGCGCGGGTTCGGACGGCACGAGCTTCGCCGA 1000

G R A K A F G A G A D G T S F A E  
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
G A G V L I V E R L S D A E R N  
GTCACACCGTCTGGCGGTGCTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100  
5 G H T V L A V V R G S A V N Q D G  
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTGCGCAGGAGCGGGTGAT 1150  
A S N G L S A P N G P S Q E R V I  
CCGGCAGGCCCCTGGCCAACGCCGGGCTCACCCCGGCGACGTGGACGCCG 1200  
R Q A L A N A G L T P A D V D A  
10 TCGAGGCCACCGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G T G T R L G D P I E A Q  
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300  
A V L A T Y G E R A T P L L L G  
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCTCCGGCGTCCGCCG 1350  
15 S L K S N I G H A Q A A S G V A  
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
G I I K M V Q A L R H G E L P P T  
CTGCACGCCGACGAGCCGTCGCCGCACGTGCGACTGGACGGCCGGCGCCGT 1450  
L H A D E P S P H V D W T A G A V  
20 CGAAGTCTGCTGACGTGCGCCCGGCGGTGGCCCGAGACCGACCGGCCACGGC 1500  
E L L T S A R P W P E T D R P R  
GTGCCGCCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCACGTCATC 1550  
R A A V S S F G V S G T N A H V I  
CTGGAGGCCGGACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGA 1600  
25 L E A G P V T E T P A A S P S G D  
CCTTCCCCTGCTGGTGTGCGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650  
L P L L V S A R S P E A L D E Q  
TCCGCCGACTGCGGCCTACCTGGACACACCCCGGACGTGACCGGGTG 1700  
I R R L R A Y L D T T P D V D R V  
30 GCCGTGGCACAGACGCTGGCCCGGCGCACACTTCGCCCACCGCGCCGT 1750  
A V A Q T L A R R T H F A H R A V  
GCTGCTCGGTGACACCGTCATCACACACCCCGCGGACCGGCCGACG 1800  
L L G D T V I T P P A D R P D  
AACTCGTCTTCTGCTACTCCGGCCAGGGCACCCAGCATCCCGCATGGGC 1850  
35 E L V F V Y S G Q G T Q H P A M G  
GAGCAGCTAGCCGATTCTGTCGTTGGTGTTCGCCGAGCGGATGGCCGAGTG 1900  
E Q L A D S S V V F A E R M A E C  
TGCGGCGGCGTTGCGCGAGTTCTGTTGACTGGGATCTGTTACGGTTCTGG 1950  
A A A L R E F V D W D L F T V L  
40 ATGATCCGCGGTGGTGACCGGTTGATGTGGTCCAGCCCGCTTCTGG 2000  
D D P A V V D R V D V V Q P A S W  
GCGATGATGGTTTCCCTGGCCGCGGTGTGGCAGGCGCGCGGTGTGCGGCC 2050  
A M M V S L A A V W Q A A G V R P  
GGATGCGGTGATCGGCCATTGCGAGGGTGAGATCGCCGCGAGCTTGTGTGG 2100  
45 D A V I G H S Q G E I A A A C V  
CGGGTGGGTGCTACTACGCGATGCCGCCCGGATCGTGACCTTGGCGAGC 2150  
A G A V S L R D A A R I V T L R S  
CAGGCGATCGCCCGGGGCTGGCGGGCGGGGCGCGATGGCATCCGTCGC 2200  
Q A I A R G L A G R G A M A S V A  
50 CCTGCCCGCGCAGGATGTGAGCTGGTTCGACGGGGCCTGGATCGCCGCC 2250  
L P A Q D V E L V D G A W I A A  
ACAACGGGCCCCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTGAC 2300  
H N G P A S T V I A G T P E A V D  
CATGTCCTCACCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCAC 2350  
55 H V L T A H E A Q G V R V R R I T  
CGTCGACTATGCCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAAC 2400  
V D Y A S H T P H V E L I R D E  
TACTCGACATCACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGG 2450  
L L D I T S D S S S Q T P L V P W  
60 CTGTGACCGTGGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTA 2500  
L S T V D G T W V D S P L D G E Y  
CTGGTACCGGAACCTGCGTGAACCGGTGCGTTTCCACCCCGCGTCAGCC 2550  
W Y R N L R E P V G F H P A V S  
AGTTGCAGGCCAGGGCGACACCGTGTTCTGTCGAGGTGAGCGCCAGCCG 2600

Q L Q A Q G D T V F V E V S A S P  
GTGTTGTTGTCAGGCGATGGACGACGATGTCGTCACGGTTGCCACGCTGCG 2650  
V L L Q A M D D D V V T V A T L R  
TCGTGACGACGGCGACGCCACCCGGATGCTCACC GCCCTGGCACAGGCCT 2700  
5 R D D G D A T R M L T A L A Q A  
ATGTCCACGGCGTCACCGTCGACTGGCCCGCCATCCTCGGCACCACCACA 2750  
Y V H G V T V D W P A I L G T T T  
ACCCGGGTACTGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTG 2800  
T R V L D L P T Y A F Q H Q R Y W  
10 GCTCGAGTCGGGACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGTGG 2850  
L E S A R P A A S D A G H P V L  
GCTCCGGTATCGCCCTCGCCGGGTCGCGGGGCCGGGTGTTACGGGTTCC 2900  
G S G I A L A G S P G R V F T G S  
GTGCCGACCGGTGCGGACCGCGCGGTGTTGTCGCGGAGCTGGCGCTGGC 2950  
15 V P T G A D R A V F V A E L A L A  
CGCCGCGGACGCGGTGACTGCGCCACGGTCGAGCGGCTCGACATCGCCT 3000  
A A D A V D C A T V E R L D I A  
CCGTGCCCCGCGCGGCCATGGCCGACGACCGTACAGACCTGGGTC 3050  
S V P G R P G H G R T T V Q T W V  
20 GACGAGCCGGCGGACGACGGCCGGCGCGGTTCACCGTGCACACCCGCAC 3100  
D E P A D D G R R R F T V H T R T  
CGGCGACGCCCCGTGGACGCTGCACGCCGAGGGGGTGTGCGCCCCCATG 3150  
G D A P W T L H A E G V L R P H  
GCACGGCCCTGCCGATGCGGCCGACGCCGAGTGGCCCCCACC GGCGCG 3200  
25 G T A L P D A A D A E W P P P G A  
GTGCCCGCGGACGGGCTGCCGGGTGTGTGGCGCCGGGGGACCAGGTCTT 3250  
V P A D G L P G V W R R G D Q V F  
CGCCGAGGCCGAGGTGGACGGACCGGACGGTTTCGTGGTGCACCCCGACC 3300  
A E A E V D G P D G F V V H P D  
30 TGCTGACGCGGTCTTCTCCGCGGTGCGCGACGGAAGCCGCCAGCCGGCC 3350  
L L D A V F S A V G D G S R Q P A  
GGATGGCGGACCTGACGGTGCACGCGTCGGACGCCACCGTACTGCGCGC 3400  
G W R D L T V H A S D A T V L R A  
CTGCCTCACCCGGCGCACCGACGGAGCCATGGGATTCGCCGCCTTCGACG 3450  
35 C L T R R T D G A M G F A A F D  
GCGCCGGCCTGCCGGTACTCACCGCGGAGGCGGTGACGCTGCGGGAGGTG 3500  
G A G L P V L T A E A V T L R E V  
GCGTCACCGTCCGGCTCCGAGGAGTCGGACGGCCTGCACCGGTTGGAGTG 3550  
A S P S G S E E S D G L H R L E W  
40 GCTCGCGGTGCGCGAGGCGGTCTACGACGGTGACCTGCCCGAGGGACATG 3600  
L A V A E A V Y D G D L P E G H  
TCCTGATCACCGCCGCCACCCCGACGACCCCGAGGACATACCCACCCGC 3650  
V L I T A A H P D D P E D I P T R  
GCCACACCCGCGCCACCCGCTCCTGACCGCCTGCAACACCACCTCAC 3700  
45 A H T R A T R V L T A L Q H H L T  
CACCACCGACCACACCTCATCGTCCACACCACCGACCCCGCCGGCG 3750  
T T D H T L I V H T T T D P A G  
CCACCGTCACCGGCTCACCGCACCGCCAGAACGAACACCCCCACCGC 3800  
A T V T G L T R T A Q N E H P H R  
50 ATCCGCTCATCGAAACCGACACCCCCACACCCCTCCCCCTGGCCCA 3850  
I R L I E T D H P H T P L P L A Q  
ACTCGCCACCTCGACCACCCCCACCTCCGCTCACCACACACCTCC 3900  
L A T L D H P H L R L T H H T L  
ACCACCCACCTACCCCTCCACACCACCCACCCACCCACCCACC 3950  
55 H H P H L T P L H T T T P P T T T  
CCCCCAACCCGAACAGCCATCATCATACCGGCGGCTCCGGCACCT 4000  
P L N P E H A I I I T G G S G T L  
CGCCGGCATCCTCGCCCGCCACCTGAACACCCCCACCTACCTCCTCT 4050  
A G I L A R H L N H P H T Y L L  
60 CCCGACCCACCCCGGACGCCACCCCGGACCCACCTCCCTGCGAC 4100  
S R T P P P D A T P G T H L P C D  
GTCGGGACCCCACTCGCCACACCTCACCCACATCCCCAACC 4150  
V G D P H Q L A T T L T H I P Q P  
CCTCACCGCATCTTCCACACCGCCGCCACCTCGACGACGGCATCCTCC 4200

```

      L T A I F H T A A T L D D G I L
ACGCCCTCACCCCGACCGCTCACCACCGTCTCCACCCCAAAGCCAAC 4250
H A L T P D R L T T V L H P K A N
GCCGCCTGGCACCCTGCACCACCTCACCCAAAACCAACCCCTCACCCACTT 4300
5  A A W H L H H L T Q N Q P L T H F
CGTCTCTACTCCAGCGCCGCGCGTCTCGGCAGCCCCGGACAAGGAA 4350
V L Y S S A A A V L G S P G Q G
ACTACGCGCGCCCAACGCCTTCCTCGACGCCCTCGCCACCCACCGCCAC 4400
N Y A A A N A F L D A L A T H R H
10 ACCCTCGGCCAACCCGCCACCTCCATCGCCTGGGGCATGTGGCACACCAC 4450
T I G Q P A T S I A W G M W H T T
CAGCACCTCACCGACAACCTCGACGACGCCGACCGGGACCGCATCCGCC 4500
S T L T G Q L D D A D R D R I R
GCGGCGGTTTCTCCGATCACGGACGACGAGGGCATGGGGATGCAT
15 R G G F L P I T D D E G

```

Phage KC515 DNA was prepared using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of confluent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *Bam*HI and *Pst*I, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes *Bgl*II and *Nsi*I and ligated into the compatible *Bam*HI and *Pst*I sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of *Streptomyces lividans* TK24 using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual edited by D. Hopwood *et al.* and overlaid with TK24 spores. After 16-24 hr, the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood *et al.*, *supra*). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya *et al.* (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

*Streptomyces hygroscopicus* ATCC 14891 (see US Patent No. 3,244,592, issued 5 Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage ( $1 \times 10^8$  of each), and incubating on R2YE agar (Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D.

Hopwood *et al.*) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 µg/ml) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain  
5 thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by replica plating onto thiostrepton  
10 containing agar medium. The PCR was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS/AT junction or the AT/DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains,  
15 followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

### Example 2

#### Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

20 The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference; *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894,366,  
25 incorporated herein by reference; *S. sp.* MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S. sp.* MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem.* 256: 528-534, and Motamedi *et al.*,  
1997, "Structural organization of a multifunctional polyketide synthase involved in the  
30 biosynthesis of the macrolide immunosuppressant FK-506," *Eur. J. Biochem.* 244: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces sp.* MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant  
35 gene clusters of the present invention differ from the naturally occurring gene clusters in

that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

- 5 The naturally occurring module 8 sequence for the MA6548 strain is shown below, followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

```

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
  M R L Y E A A R R T G S P V V V
GCGGCCGCECTCGACGACGCGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
  A A A L D D A P D V P L L R G L R
10 GCGTACGACCGTCCGGCGGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
  R T T V R R A A V R E R S L A D
GCTCGCCGTGCTGCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTTCG 200
  R S P C C P T T S A P T P P S R S
15 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
  S W N S T A T V L G H L G A E D I
CCC GCGACGACGACGTTC AAGGA ACTCGGCATCGACTCGCTCACC GCG 300
  P A T T T F K E L G I D S L T A
TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
20 V Q L R N A L T T A T G V R L N A
ACAGCGGTCTTCGACTTTCGACGCCGCGCGCGCTCGCCGCGAGACTCGG 400
  T A V F D F P T P R A L A A R L G
CGACGAGCTGGCCGGTACCCGCGCGCCCGTCCGGGCGGACCGCGGCCA 450
  D E L A G T R A P V A A R T A A
25 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
  T A A A H D E P L A I V G M A C R
CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
  L P G G V A S P Q E L W R L V A S
CGGCACCGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
30 G T D A I T E F P A D R G W D V
ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
  D A L Y D P D P D A I G K T F V R
CACGGCGGCTTCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
  H G G F L D G A T G F D A A F F G
35 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
  I S P R E A L A M D P Q R V L
TGGAGACGTCTCTGGGAGGCGTTTCAAAGCGCGGCATCACCCCGGACGCG 800
  L E T S W E A F E S A G I T P D A
GCGCGGGGACGCGACACCGGCGTGTTTCATCGGCGGTTCTCCTACGGGTA 850
40 A R G S D T G V F I G A F S Y G Y
CGGCACGCGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGCA 900
  G T G A D T N G F G A T G S Q T
GCGTGCTCTCCGGCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
  S V L S G R L S Y F Y G L E G P S
45 GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000
  V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050
  G Q S L R S G E C S L A L V G G
TCACGGTGATGGCGTCGCCCGGGGATTTCGTGAGTTCTCCCGGACGCGC 1100
50 V T V M A S P G G F V E F S R Q R
GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
  G L A P D G R A K A F G A G A D G
TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTCGAGCGGCTCTCCG 1200
  T S F A E G A G A L V V E R L S
55 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250
  D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGGTGCAACGGTCTGTGCGGCGCGAAGCGCCCTC 1300
  A N S D G A S N G L S A P N G P S
CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAACTCACCCCGG 1350

```



Q E R V I H Q A L A N A K L T P  
 CCGATGTGACGCGGTGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400  
 A D V D A V E A H G T G T R L G D  
 5 CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
 P I E A Q A L L A T Y G Q D R A T  
 GCCCCTGCTGCTCGGCTCGTGAAGTCGAACATCGGGCACGCCAGGCCG 1500  
 P L L L G S L K S N I G H A Q A  
 CGTCAGGGGTGCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
 A S G V A G I I K M V Q A I R H G  
 10 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCCGACGTGACTG 1600  
 E L P P T L H A D E P S P H V D W  
 GACGGCCGGTCCGCTCGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGA 1650  
 T A G A V E L L T S A R P W P G  
 CCGGTGCGCCGCGCGCGCTGCCGTCTCGTCTCGGCGTGAGCGGCACG 1700  
 15 T G R P R R A A V S S F G V S G T  
 AACGCCCACATCATCTTGAGGCAGGACCGGTCAAAACGGGACCGGTGCA 1750  
 N A H I I L E A G P V K T G P V E  
 GGCAGGAGCGATCGAGGCAGGACCGGTCAAGTAGGACCGGTGAGGCTG 1800  
 A G A I E A G P V E V G P V E A  
 20 GACCGCTCCCCGCGCGCGCGCGCTCAGCACCGGGCGAAGACCTTCCGCTG 1850  
 G P L P A A P P S A P G E D L P L  
 CTCGTGTGCGGCGGTTCCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
 L V S A R S P E A L D E Q I G R L  
 GCGCGCCTATCTCGACACCGGCCCGGGCGTGCACCGGGCGGCCGTGGCGC 1950  
 25 R A Y L D T G P G V D R A A V A  
 AGACACTGGCCCGCGGTACGCACTTCACCCACCGGGCCGTACTGCTCGGG 2000  
 Q T L A R R T H F T H R A V L L G  
 GACACCGTCATCGGCGCTCCCCCGCGGACCAGGCCGACGAACCTCGTCTT 2050  
 D T V I G A P P A D Q A D E L V F  
 30 CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100  
 V Y S G Q G T Q H P A M G E Q L  
 CGGCCGCTTCCCCGTGTTCCCGATGCCTGGCAGCAGCGCTCCGACGG 2150  
 A A A F P V F A D A W H D A L R R  
 CTCGACGACCCCGACCCGACGACCCACACGGAGCCAGCACACGCTCTT 2200  
 35 L D D P D P H D P T R S Q H T L F  
 CGCCACACAGGCGGCGTTCACCGCCCTCCTGAGGTCTGGGACATCACGC 2250  
 A H Q A A F T A L L R S W D I T  
 CGCAGCCCGTCATCGGCCACTCGCTCGGCGAGATCACCGCCGCGTACGCC 2300  
 P H A V I G H S L G E I T A A Y A  
 40 GCCGGGATCCTGTGCTCGACGACGCTGCACCCTGATCACACGCGTGC 2350  
 A G I L S L D D A C T L I T T R A  
 CCGCCTCATGCACACGCTTCGCGCGCCGCGCCATGGTCACCGTGCTGA 2400  
 R L M H T L P P P G A M V T V L  
 CCAGCGAGGAGGAGGCCCGTCAGGCGCTGCGGCCGGGCGTGAGATCGCC 2450  
 45 T S E E E A R Q A L R P G V E I A  
 GCGTCTTCCGCCCGCACTCCGTGCTGCTCTCGGGCGACGAGGACGCCGT 2500  
 A V F G P H S V V L S G D E D A V  
 GCTCGACGTGCACACGCGCTCGGCATCCACCACCGTCTGCCCGCGCCGC 2550  
 L D V A Q R L G I H H R L P A P  
 50 ACGCGGGCCACTCCGCGCACATGGAACCCGTGGCCGCCGAGCTGCTCGCC 2600  
 H A G H S A H M E P V A A E L L A  
 ACCACTCGCGAGTCCGTTACGACCGGCCACACCGCCATCCCGAACGA 2650  
 T T R E L R Y D R P H T A I P N D  
 CCCCACACCGCCGAGTACTGGGCGGAGCAGGTCCGCAACCCCGTGCTGT 2700  
 55 P T T A E Y W A E Q V R N P V L  
 TCCACGCCCACACCCAGCGGTACCCCGACCGGTGTTGTCGAGATCGGC 2750  
 F H A H T Q R Y P D A V F V E I G  
 CCCGGCCAGGACCTCTACCGCTGGTGCAGGCATCGCCCTGCAGAACGG 2800  
 P G Q D L S P L V D G I A L Q N G  
 60 CACGGCGGACGAGGTGCACGCGTGCACACCGCGCTCGCCCGCCTCTTCA 2850  
 T A D E V H A L H T A L A R L F  
 CACGCGGCGCCACGCTCGACTGGTCCCGCATCTCGGCGGTGCTTCGCGG 2900  
 T R G A T L D W S R I L G G A S R  
 CACGACCCTGACGTCCCCTCGTACGCGTTCAGCGCGTCCCTACTGGAT 2950

H D P D V P S Y A F Q R R P Y W I  
CGAGTCGGCTCCCCGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCA 3000  
E S A P P A T A D S G H P V L G  
5 CCGGAGTCGCGTTCGCGGGTTCGCGGGCGGGTGTTCACGGTCCCGTG 3050  
T G V A V A G S P G R V F T G P V  
CCCGCCGGTTCGGACCGCGCGGTGTTCATCGCCGAACGGCGCTCGCCGC 3100  
P A G A D R A V F I A E L A L A A  
CGCCGACGCCACCGACTGCGCCACGGTTCGAACAGCTCGACGTCACCTCCG 3150  
A D A T D C A T V E Q L D V T S  
10 TGCCCGGGCGATCCGCGCGGGCAGGGCCACCGCGCAGACCTGGGTTCGAT 3200  
V P G G S A R G R A T A Q T W V D  
GAACCCGCGCGGACGGGCGGGCGCGCTTCACCGTCCACACCCGCGTCGG 3250  
E P A A D G R R R F T V H T R V G  
CGACGCCCCGTGGACGCTGCACGCCGAGGGGGTCTCCGCCCCGGCCGCG 3300  
15 D A P W T L H A E G V L R P G R  
TGCCCGAGCCGAAGCCGTTCGACACCGCTGGCCCCCGCGGGCGCGGTG 3350  
V P Q P E A V D T A W P P P G A V  
CCCGCGGACGGGCTGCCGCGGGCGTGGCGACGCGCGGACCAGGTCTTCGT 3400  
20 P A D G L P G A W R R A D Q V F V  
CGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGC 3450  
E A E V D S P D G F V A H P D L  
TCGACGCGGTCTTCTCCGCGGTTCGGCGACGGGAGCCGCCAGCCGACCGGA 3500  
L D A V F S A V G D G S R Q P T G  
25 TGGCGCGACCTCGCGGTGCACGCGTTCGGACGCCACCGTGTCTGCGCGCCTG 3550  
W R D L A V H A S D A T V L R A C  
CCTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCGCCCTTCGACGGTG 3600  
L T R R D S G V V E L A A F D G  
CCGGAATGCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTTCGCG 3650  
30 A G M P V L T A E S V T L G E V A  
TCGGCAGGCGGATCCGACGAGTTCGGACGGTCTGCTTCGGCTTGAGTGGTT 3700  
S A G G S D E S D G L L R L E W L  
GCCGGTGGCGGAGGCCCACTACGACGGTCCCGACGAGCTGCCCCGAGGGCT 3750  
P V A E A H Y D G A D E L P E G  
35 ACACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCCAAC 3800  
Y T L I T A T H P D D P D D P T N  
CCCCACAACACCCACACGACCCACACACAAACCACACGCGTCTCTAC 3850  
P H N T P T R T H T Q T T R V L T  
CGCCCTCCAACACCACCTCATCACCAACCAACACCCCTCATCGTCCACA 3900  
A L Q H H L I T T N H T L I V H  
40 CCACCACCGACCCCCAGGCGCCGCGTACCGGCCTCACCCGACCGCA 3950  
T T T D P P G A A V T G L T R T A  
CAAAACGAACACCCCGCGCCGATCCACCTCATCGAAACCCACACCCCCA 4000  
Q N E H P G R I H L I E T H H P H  
45 CACCCCACTCCCCCTCACCCAACTACCAACCCCTCCACCAACCCACCTAC 4050  
T P L P L T Q L T T L H Q P H L  
GCCTCACCAACAACACCTCCACACCCCCACCTACCCCATCACCAAC 4100  
R L T N N T L H T P H L T P I T T  
CACCACAACACCACAACCAACCCCAACACCCCAACCCCTCAACCCAA 4150  
H H N T T T T T P N T P P L N P N  
50 CCACGCCATCCTCATCACCGGCGGCTCCGGCACCCCTCGCGGCATCCTCG 4200  
H A I L I T G G S G T L A G I L  
CCCGCCACCTCAACACCCCAACACCTACCTCTCTCCGACACCAACCA 4250  
A R H L N H P H T Y L L S R T P P  
55 CCCCCACACACCCGGCACCCACATCCCCTGCGACCTACCGACCCAC 4300  
P P T T P G T H I P C D L T D P T  
CCAAATCACCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCT 4350  
Q I T Q A L T H I P Q P L T G I  
TCCACACCGCGCCACCCCTCGACGACGCCACCCCTACCAACCTCACCCCC 4400  
F H T A A T L D D A T L T N L T P  
60 CAACACCTCACCAACCCCTCAACCCAAAGCCGACGCGCCTGGCACCT 4450  
Q H L T T T L Q P K A D A A W H L  
CCACCACACACCCAAACCAACCCCTCACCACTTCGTCTCTACTCCA 4500  
H H H T Q N Q P L T H F V L Y S  
GCGCCGCGCCACCCCTCGGCAGCCCCGGCCAAGCCAACCTACGCGCGCGCC 4550

S A A A T L G S P G Q A N Y A A A  
 AACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACC 4600  
 N A F L D A L A T H R H T Q G Q P  
 CGCCACCACCATCGCCTGGGGCATGTGGCACACCACCACTCACCA 4650  
 5 A T T I A W G M W H T T T T L T  
 GCCAACTCACCACGACGACCGCGACCGCATCCGCCGCGGCGGCTTCCTG 4700  
 S Q L T D S D R D R I R R G G F L  
 CCGATCTCGGACGACGAGGGCATGC  
 10 P I S D D E G M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 M R L Y E A A R R T G S P V V V  
 15 GCGGCCGCGCTCGACGACGCGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 A A A L D D A P D V P L L R G L R  
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
 R T T V R R A A V R E R S L A D  
 20 GCTCGCCGTGCTGCCCGACGACGAGCGCGCCGACGCTCCCTCGCGTTG 200  
 R S P C C P T T S A P T P P S R S  
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 CCGGCGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300  
 P A T T T F K E L G I D S L T A  
 25 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGCGTACGCCTCAACGCC 350  
 V Q L R N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400  
 T A V F D F P T P R A L A A R L G  
 CGACGAGCTGGCCGGTACCCGCGCGCCGTCGCGGCCCGGACCGCGGCCA 450  
 30 D E L A G T R A P V A A R T A A  
 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 L P G C V A S P Q E L W R L V A S  
 35 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600  
 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 D A L Y D P D P D A I G K T F V R  
 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700  
 40 H G G F L D G A T G F D A A F F G  
 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGAGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q Q R V L  
 TGGAGACGTCTGGGAGGCGTTTCAAAGCGCGGGCATCACCCCGGACGCG 800  
 L E T S W E A F E S A G I T P D A  
 45 GCGCGGGGACGCGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850  
 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGAGACCA 900  
 G T G A D T N G F G A T G S Q T  
 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 50 S V L S G R L S Y F Y G L E G P S  
 GTCACGGTGCACACCGCCTGCTCGTCTGCTACTGGTCCGCTGCACAGGC 1000  
 V T V D T A C S S S L V A L H Q A  
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTGGGCGGTG 1050  
 G Q S L R S G E C S L A L V G G  
 55 TCACGGTGATGGCGTCCGCCGGGATTTCGTGAGTTCTCCCGGACGCGC 1100  
 V T V M A S P G G F V E F S R Q R  
 GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150  
 G L A P D G R A K A F G A G A D G  
 TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200  
 60 T S F A E G A G A L V V E R L S  
 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250  
 D A E R H G H T V L A L V R G S A

GCTAACTCCGACGGCGCGTCGAACGGTCTGTGGCGCCGAACGGCCCCCTC 1300  
A N S D G A S N G L S A P N G P S  
CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCG 1350  
Q E R V I H Q A L A N A K L T P  
5 CCGATGTCGACGCGGTGAGGCGCACGGCACCCGCGCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D R A T  
GCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500  
10 P L L L G S L K S N I G H A Q A  
CGTCAGGGGTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCGCACGTGACTG 1600  
E L P P T L H A D E P S P H V D W  
15 GACGGCCGGTGCCGTCGAGCTCCTGACGTGCGCCCCGGCCGTGGCCGGGA 1650  
T A G A V E L L T S A R P W P G  
CCGGTCGCCCCAGGCGGCGGTGCTCCTTCGGGATCAGTGGCACC 1700  
T G R P R R A G V S S F G I S G T  
AACGCCACGTCATCCTGGAAGCGCACCCCCACTCAGCCTGCGGACAA 1750  
20 N A H V I L E S A P P T Q P A D N  
CGCGGTGATCGAGCGGGCACCGGAGTGGGTGCCGTTGGTGATTTCGGCCA 1800  
A V I E R A P E W V P L V I S A  
GGACCCAGTCGGCTTTGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTG 1850  
R T Q S A L T E H E G R L R A Y L  
25 GCGGCGTCGCCCCGGGTGGATATGCGGGCTGTGGCATCGACGCTGGCGAT 1900  
A A S P G V D M R A V A S T L A M  
GACACGGTCGGTGTTGAGCACCGTGCCGTGCTGCTGGGAGATGACACCG 1950  
T R S V F E H R A V L L G D D T  
TCACCGGCACCCTGTGTCTGACCCTCGGGCGGTGTTGCTCTTCCCGGA 2000  
30 V T G T A V S D P R A V F V F P G  
CAGGGGTGCGCAGCGTGCTGGCATGGGTGAGGAACTGGCCGCGCGTTCCC 2050  
Q G S Q R A G M G E E L A A A F P  
CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCG 2100  
V F A R I H Q Q V W D L L D V P  
35 ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTGCAATG 2150  
D L E V N E T G Y A Q P A L F A M  
CAGGTGGCTCTGTTGCGGCTGCTGGAATCGTGGGGTGTACGACCGGACGC 2200  
Q V A L F G L L E S W G V R P D A  
GGTGATCGGCCATTGCGGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGG 2250  
40 V I G H S V G E L A A A Y V S G  
TGTGGTGGTGGAGGATGCCTGCACTTTGGTGTCGGCGCGGGCTCGTCTG 2300  
V W S L E D A C T L V S A R A R L  
ATGAGGCTCTGCCCCGCGGTGGGGTGGTGGTGGTGTGCTGCTCCCGGTCTCGGA 2350  
M Q A L P A G G V M V A V P V S E  
45 GGATGAGGCCCGGGCGGTGCTGGGTGAGGGTGTGGAGATCGCCGCGGTCA 2400  
D E A R A V L G E G V E I A A V  
ACGGCCCCGTGCTGGGTGGTTCTCTCCGGTGATGAGGCCGCGGTGCTGCAG 2450  
N G P S S V V L S G D E A A V L Q  
GCCGCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGACCAGCCACGCGTT 2500  
50 A G E L G K W T R L A T S H A F  
CCATTCCGCCCCGATGGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCG 2550  
H S A R M E P M L E E F R A V A  
AAGGCCTGACCTACCGGACGCGCAGGTCTCCATGGCCGTTGGTGATCAG 2600  
E G L T Y R T P Q V S M A V G D Q  
55 GTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGGACACGGTCCGTT 2650  
V T T A E Y W R Q V R D T V R F  
CGGCGAGCGGTGGCCTCGTACGAGGACGCGGTGTTGCTCGAGCTGGGTG 2700  
G E Q V A S Y E D A V F V E L G  
CCGACCGGTCACTGGCCCGCTGGTGCACGGTGTGCGGATGCTGCACGGC 2750  
60 A D R S L A R L V D G V A M L H G  
GACCACGAAATCCAGGCCGCGATCGGGCCCTGGCCACCTGTATGTCAA 2800  
D H E I Q A A I G A L A H L Y V N  
CGGCGTCACGGTCGACTGGCCCGGCTCCTGGGCGATGCTCCGGCAACAC 2850  
G V T V D W P A L L G D A P A T

GGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCAGCGCTACTGGCTC 2900  
R V L D L P T Y A F Q H Q R Y W L  
GAGTCGGCTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCCTCGGCAC 2950  
E S A P P A T A D S G H P V L G T  
5 CGGAGTCGCGGTGCGCGGGTGGCGGGCGGGTGTTCACGGGTCCCGTGC 3000  
G V A V A G S P G R V F T G P V  
CCGCGGTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCC 3050  
P A G A D R A V F I A E L A L A A  
10 GCCGACGCCACCGACTGCGCCACGGTCTGAACAGCTCGACGTCACCTCCGT 3100  
A D A T D C A T V E Q L D V T S V  
GCCCGGCGGATCCGCGCGGCGAGGGCCACCGCGCAGACCTGGGTGCGATG 3150  
P G C S A R G R A T A Q T W V D  
AACCCGCGCGGACGGGCGCGCTTACCGTCCACACCCGCGTCCGGC 3200  
E P A A D G R R R F T V H T R V G  
15 GACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGGCGCGT 3250  
D A P W T L H A E G V L R P G R V  
GCCCGAGCCCGAAGCCGTGACACCGCCTGGCCCCCGCGGGCGCGGTGC 3300  
P Q G P E A V D T A W P P P G A V  
CCGCGGACCGCTGCCCCGGGCGTGGCGACGCGCGGACCAGGTCTTCGTC 3350  
20 P A D G L P G A W R R A D Q V F V  
GAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCT 3400  
E A E V D S P D G F V A H P D L L  
CGACGCGGTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGAT 3450  
D A V F S A V G D G S R Q P T G  
25 GCGCGACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGC 3500  
W R D L A V H A S D A T V L R A C  
CTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCGCCTTCGACGGTGC 3550  
L T R R D S G V V E L A A F D G A  
CGGAATGCCGGTGTCTACCGCGGAGTCCGGTACGCTGGGCGAGGTGCGGT 3600  
30 G M P V L T A E S V T L G E V A  
CGGCAGGCGGATCCGACGAGTCCGACGGTCTGCTTCGGCTTGAGTGGTTG 3650  
S A G S D E S D G L R L E W L  
CCGGTGGCGGAGGCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTA 3700  
P V A E A H Y D G A D E L P E G Y  
35 CACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCACCAACC 3750  
T L I T A T H P D D P D D P T N  
CCCACAACACACCCACACGACCCACACACAAACCACACGCGTCCCTCACC 3800  
P H N T P T R T H T Q T T R V L T  
GCCCTCCAACACCACTCATCACCAACCAACCCCTCATCGTCCACAC 3850  
40 A L Q H H L I T T N H T L I V H T  
CACCACCGACCCCCAGGCGCGCGCTCACC GGCTCACC CGCACCGCAC 3900  
T T D P P G A A V T G L T R T A  
AAAACGAACACCCCGCGCATCCACCTCATCGAAACCCACCAACCCAC 3950  
Q N E H P G R I H L I E T H H P H  
45 ACCCACTCCCCCTACCCAACCTCACCACCTCCACCAACCCACCTACG 4000  
T P L T Q L T T L H Q P H L R  
CCTCACCAACAACACCTCCACACCCCCACCTCACC CCATCACCACCC 4050  
L T N N T L H T P H L T P I T T  
ACCACAACACCACACAAACCACCCCAACACCCACCCCTCAACCCCAAC 4100  
50 H H N T T T T T P N T P P L N P N  
CACGCCATCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGC 4150  
H A I L I T G G S G T L A G I L A  
CGGCACCTCAACCACCCCAACCTACCTCCTCCTCCCGCACACCACAC 4200  
R H L N H P H T Y L L S R T P P  
55 CCCCCACCAACCCCGGACCCACATCCCCTGCGACCTCACC GACCCAC 4250  
P P T T P G T H I P C D L T D P T  
CAAATCACC CAAGCCCTCACCACATACCACAACCCCTCACC GGCATCTT 4300  
Q I T Q A L T H I P Q P L T G I F  
CCACACCGCGCCACCTCGACGACGCCACCTCACC AACCTCACC CCCC 4350  
60 H T L A A T D D A T L T N L T P  
AACACCTCACCACACCCCTCCAACCCAAAGCCGACGCGCGCTGGCACCTC 4400  
Q H L T T T L Q P K A D A A W H L  
CACCACCAACCCAAAACCAACCCCTCACC ACTTCGTCTCTACTCCAG 4450  
H H H T Q N Q P L T H F V L Y S S

CGCCGCCGCCACCCCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCCGCCA 4500  
 A A A T L G S P G Q A N Y A A A  
 ACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCC 4550  
 N A F L D A L A T H R H T Q G Q P  
 5 GCCACCACCATCGCCTGGGGCATGTGGCACACCACCACACTCACCAG 4600  
 A T T I A W G M W H T T T T L T S  
 CCAACTCACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCCTGC 4650  
 Q L T D S D R D R I R R G G F L  
 CGATCTCGGACGACGAGGGCATGC  
 10 P I S D D E G M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 15 M R L Y E A A R R T G S P V V V  
 GCGGCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 A A A L D D A P D V P L L R G L R  
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
 R T T V R R A A V R E R S L A D  
 20 GCTCGCCGTGCTGCCCAGCAGAGCGCGCGACGCCTCCCTCGCGTTCG 200  
 R S P C P T T S A P T P P S R S  
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 CCCGGCGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300  
 25 P A T T T F K E L G I D S L T A  
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
 V Q L R N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCTCGCCGCGAGACTCGG 400  
 T A V F D F P T P R A L A A R L G  
 30 CGACGAGCTGCGCGGTACCCGCGCGCCCGTCCGGGCCCGGACCGCGGCCA 450  
 D E L A G T R A P V A A R T A A  
 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 CTGCCGGGCGGGTTCGCGTCCGACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 35 L P G G V A S P Q E L W R L V A S  
 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600  
 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 D A L Y D P D P D A I G K T F V R  
 40 CACGGCGGCTTCCTCGACGGTTCGACCGGCTTCGACGCGGCGTTCCTCGG 700  
 H G F L D G A T G F D A A F F G  
 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGAGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q Q R V L  
 TGGAGACGTCTGGGAGGCGTTCGAAAGCGCGGCATACCCCGGACGCG 800  
 45 L E T S W E A F E S A G I T P D A  
 GCGCGGGGCGAGCGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850  
 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAACGGGTTTCGGCGCGACAGGGTTCGACACCA 900  
 G T G A D T N G F G A T G S Q T  
 50 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 S V L S G R L S Y F Y G L E G P S  
 GTCACGGTGCACACCGCCTGCTCGTCTCACTGGTCGCCCTGCACCAGGC 1000  
 V T V D T A C S S S L V A L H Q A  
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050  
 55 G Q S L R S G E C S L A L V G G  
 TCACGGTGATGGCGTCGCCCCGGCGGATTCGTGAGTTCTCCCGGCAGCGC 1100  
 V T V M A S P G G F V E F S R Q R  
 GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGACGG 1150  
 G L A P D G R A K A F G A G A D G  
 60 TACGAGCTTCGCCGAGGGCGCCGGTGGCCTGGTGGTTCGAGCGGCTCTCCG 1200  
 T S F A G A L V E R L S  
 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250

D A E R H G H T V L A L V R G S A  
 GCTAACTCCGACGGCGCTCGAACGGTCTGTGCGGCGCCGAACGGCCCCTC 1300  
 A N S D G A S N G L S A P N G P S  
 5 CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350  
 Q E R V I H Q A L A N A K L T P  
 CCGATGTGACGCGGTGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400  
 A D V D A V E A H G T G T R L G D  
 CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
 P I E A Q A L L A T Y G Q D R A T  
 10 GCCCCGTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCCAGGCCG 1500  
 P L L L G S L K S N I G H A Q A  
 CGTCAGGGGTGCGCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
 A S G V A G I I K M V Q A I R H G  
 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCCGACGTCGACTG 1600  
 15 E L P P T L H A D E P S P H V D W  
 GACGGCCGGTGCCGTGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGGA 1650  
 T A G A V E L L T S A R P W P G  
 CCGGTCGCCCCTAGGCGGGCGGGCTGTCTCCTTCGGAGTCAGCGGCACC 1700  
 T G R P R R A G V S S F G V S G T  
 20 AACGCCCCACGTCATCCTGGAGAGCGCACCCCCCGCTCAGCCCCGCGGAGGA 1750  
 N A H V I L E S A P P A Q P A E E  
 GCGCGAGCCTGTTGAGACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGG 1800  
 A Q P V E T P V V A S D V L P L  
 TGATATCGGCCAAGACCCAGCCCCGCCCTGACCGAACACGAAGACCGGCTG 1850  
 25 V I S A K T Q P A L T E H E D R L  
 CGCGCCTACCTGGCGGCTCGCCCGGGCGGATATACGGGCTGTGGCATC 1900  
 R A Y L A A S P G A D I R A V A S  
 GACGCTGGCGGTGACACGGTCCGGTGTTCGAGCACCGCGCCGTACTCCTTG 1950  
 T L A V T R S V F E H R A V L L  
 30 GAGATGACACCGTCAACCGGCACCGCGGTGACCGACCCAGGATCGTGTTT 2000  
 G D D T V T G T A V T D P R I V F  
 GTCTTTCCCGGCGAGGGGTGGCAGTGGCTGGGGATGGGCAGTGCAGTGGC 2050  
 V F P G Q W L G M G S A L R  
 CGATTCTGTCGGTGGTGTTCGCCCAGCGGATGGCCGAGTGTGCGGCGGCGT 2100  
 35 D S S V V F A E R M A E C A A A  
 TGCGCGAGTTCTGTTGACTGGGATCTGTTACGGTCTTGGATGATCCGGCG 2150  
 L R E F V D W D L F T V L D D P A  
 GTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGT 2200  
 V V D R V D V V Q P A S W A M M V  
 40 TTCCCTCGCCGCGGTGTGGCAGGCGGCGGTGTGCGGCCGGATGCGGTGA 2250  
 S L A A V W Q A A G V R P D A V  
 TCGGCCATTTCGAGGTGAGATCGCCGACGCTTGTGTGGCGGGTGGCGGTG 2300  
 I G H S Q G E I A A A C V A G A V  
 TCACTACGCGATGCCGCCCGGATCGTGACCTTGCGCAGCCAGGCGATCGC 2350  
 45 S L R D A A R I V T L R S Q A I A  
 CCGGGGCTGGCGGGCGGGCGGATGGCATCCGTGCGCCCTGCCCCGCGC 2400  
 R G L A G R G A M A S V A L P A  
 AGGATGTGAGCTGGTTCGACGGGGCCTGGATCGCCGCCACAACGGGGCCC 2450  
 Q D V E L V D G A W I A A H N G P  
 50 GCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTGACCATGTCTCTCAC 2500  
 A S T V I A G T P E A V D H V L T  
 CGCTCATGAGGCACAAGGGGTGCGGGTGGCGGATCACCGTCGACTATG 2550  
 A H E A Q G V R V R I T V D Y  
 CCTCGCACACCCCGCACGTCGAGCTGATCCGCGACGAACACTACTCGACATC 2600  
 55 A S H T P H V E L I R D E L L D I  
 ACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTGACCGT 2650  
 T S D S S S Q T P L V P W L S T V  
 GGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTACTGGTACCGGA 2700  
 D G T W V D S P L D G E Y W Y R  
 60 ACCTGCGTGAACCGGTGCGTTTCCACCCCGCGTCAGCCAGTTGCAGGCC 2750  
 N L R E P V G F H P A V S Q L Q A  
 CAGGGCGACACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGCA 2800  
 Q G D T V F V E V S A S P V L L Q  
 GGCGATGGACGACGATGTCGTACGGTTGCCACGCTGCGTCTGACGACG 2850

A M D D D V V T V A T L R R D D  
GCGACGCCACCCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGGC 2900  
G D A T R M L T A L A Q A Y V H G  
5 GTCACCGTCGACTGGCCCGCCATCCTCGGCACCACCACAACCCGGGTACT 2950  
V T V D W P A I L G T T T T R V L  
GGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGG 3000  
D L P T Y A F Q H Q R Y W L E S  
CTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCCTCGGCACCCGGAGTC 3050  
A P P A T A D S G H P V L G T G V  
10 GCCGTCGCCGGGTGCGCCGGCGGGTGTTCACGGGTCCCGTGCCCGCCGG 3100  
A V A G S P G R V F T G P V P A G  
TGCGGACCGCGGGTGTTCATCGCCGAACCTGGCGCTCGCCGCGCCGACG 3150  
A D R A V F I A E L A L A A A D  
CCACCGACTGCGCCACGGTGAACAGCTCGACGTCACCTCCGTGCCCGGC 3200  
15 A T D C A T V E Q L D V T S V P G  
GGATCCGCCCCGCGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCGC 3250  
G S A R G R A T A Q T W V D E P A  
CGCCGACGGCGCGCGCTTACCGTCCACACCCGCGTCGGCGACGCCC 3300  
A D G R R R R F T V H T R V G D A  
20 CGTGGACGCTGCACGCCGAGGGGTTCTCCGCCCCGCGCGGTGCCCCAG 3350  
P W T L H A E G V L R P G R V P Q  
CCCGAAGCCGTCGACACCGCCTGGCCCCCGCGGGCGCGGTGCCCGCGGA 3400  
P E A V D T A W P P P G A V P A D  
CGGGCTGCCCGGGCGGTGGCGACGCGCGGACCAGGTCTTCGTGAAGCCG 3450  
25 G L P G A W R A D Q V F V E A  
AAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGCG 3500  
E V D S P D G F V A H P D L L D A  
GTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGATGGCGCGA 3550  
V F S A V G D G S R Q P T G W R D  
30 CCTCGCGGTGCACGCGTCGGACGCCACCGTGCTGCGCGCCTGCCCTACCC 3600  
L A V H A S D A T V L R A C L T  
GCCGCGACAGTGGTGTCTCGTGAGCTCGCCGCTTCGACGGTGCCGGAATG 3650  
R R D S G V V E L A A F D G A G M  
CCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTGCGGTGCGGCAGG 3700  
35 P V L T A E S V T L G E V A S A G  
CGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTGG 3750  
G S D E S D G L L R L E W L P V  
CGGAGGCCCCACTACGACGGTGCCGACGAGTGGCCGAGGGCTACACCCCTC 3800  
A E A H Y D G A D E L P E G Y T L  
40 ATCACCGCCACACACCCCGACGACCCCGACGACCCCAACACCCCAACAA 3850  
I T A T H P D D P D D P T N P H N  
CACACCCACACGACCCACACACAAACCACACGCGTCTCACCGCCCTCC 3900  
T P T R T H T Q T T R V L T A L  
AACACCACTCATCACCAACCAACCACTCATCGTCCACACCAACCACC 3950  
45 Q H H L I T T N H T L I V H T T T  
GACCCCCCAGGCGCGCGCTCACCGGCTCACCCGACCCGACACAAAACGA 4000  
D P P G A A V T G L T R T A Q N E  
ACACCCCGCGCATCCACCTCATCGAAACCCACACCCCAACACCCAC 4050  
H P G R I H L I E T H H P H T P  
50 TCCCCCTCACCAACTCACCACTCCACCAACCCCACTACGCCTCACCC 4100  
L P L T Q L T T L H Q P H L R L T  
AACAACACCTCCACACCCCACTCACCCCATCACCAACCAACCAAA 4150  
N N T L H T P H L T P I T T H H N  
CACCACCAACCAACCCCAACACCCCAACCCCTCAACCCCAACCAACGCCA 4200  
55 T T T T T P N T P P L N P N H A  
TCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGCGATCCTCGCCCGCCAC 4250  
I L I T G G S G T L A G I L A R H  
CTCAACCAACCCCAACCTACCTCCTCTCCCGCACCAACCAACCCCAAC 4300  
L N H P T Y L L S R T P P P P T  
60 CACACCCGGCACCCACATCCCCTGCGACCTACCGACCCCAACCAATCA 4350  
T P G T H I P C D L T D P T Q I  
CCCAAGCCCTCACCAACATACCAACCCCTACCGGCATCTTCCACACC 4400  
T Q A L T H I P Q P L T G I F H T  
GCCGCCACCCCTCGACGACGCCACCCCTACCAACCTCACCCCAACACCT 4450



A A T L D D A T L T N L T P Q H L  
 CACCACCACCTCCAACCCAAAGCCGACGCCGCTGGCACCTCCACCACC 4500  
 T T T L Q P K A D A A W H L H H  
 ACACCCAAAACCAACCCCTCACCCTTCTGCTCTACTCCAGCGCCGCC 4550  
 5 H T Q N Q P L T H F V L Y S S A A  
 GCCACCTCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCCGCCAACGCCTT 4600  
 A T L G S P G Q A N Y A A A N A F  
 CCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACCA 4600  
 L D A L A T H R H T Q G Q P A T  
 10 CCATCGCCTGGGGCATGTGGCACACCACCACCACTCACCAGCCAACTC 4700  
 T I A W G M W H T T T T L T S Q L  
 ACCGACAGCGACCGGACCGCATCCGCCGCGGGCTTCCTGCCGATCTC 4750  
 T D S D R D R I R R G G F L P I S  
 GGACGACGAGGGCATGC  
 15 D D E G M

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 20 M R L Y E A A R R T G S P V V V  
 GCGGCGCGCTCGACGACGCGCCGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 A A A L D D A P D V P L R G L R  
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
 R T T V R R A A V R E R S L A D  
 25 GCTCGCCGTGCTGCCCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTCTG 200  
 R S P C C P T T S A P T P P S R S  
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 CCGGCGGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACC GCGG 300  
 30 P A T T T F K E L G I D S L T A  
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
 V Q L P N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCCGACGCCGCGCGCTCGCCGCGAGACTCGG 400  
 T A V F D F P T P R A L A A R L G  
 35 CGACGAGCTGGCCGTACCCGCGCGCCCGTCCGCGGCCCGGACCGCGGCCA 450  
 D E L A G T R A P V A A R T A A  
 CCGCGCGCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 CTGCCGCGCGGGTCCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 40 L P G G V A S P Q E L W R L V A S  
 CGGACCCGACGCCATCAGGAGTTCCCGCGGACCGCGGTGGGACGTGG 600  
 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 D A L Y D P D P D A I G K T F V R  
 45 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700  
 H G G F L D G A T G F D A A F F G  
 GATAGCCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q Q R V L  
 TGGAGACGTCTCGGGAGGCGTTCCGAAAAGCGCGGGCATCACCCCGGACGCG 800  
 50 L E T S W E A F E S A G I T P D A  
 GCGCGGGGCGGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850  
 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGAGACCA 900  
 G T G A D T N G F G A T G S Q T  
 55 GCGTGCTCTCCGGCCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 S V L S G R L S Y F Y G L E G P S  
 GTCACGGTCGACACCGCCTGCTCGTCGTCCTGGTTCGCCCTGCACCAGGC 1000  
 V T V D T A C S S S L V A L H Q A  
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050  
 60 G Q S L R S G E C S L A L V G G  
 TCACGGTGATGGCGTCGCCGCGGATTCGTGAGTTCTCCCGGACGCGC 1100  
 V T V M A S P G G F V E F S R Q R

GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150  
G L A P D G R A K A F G A G A D G  
TACGAGCTTCGCCGAGGGCGCCGGTGCCTGGTGGTTCGAGCGGCTCTCCG 1200  
T S F A E G A G A L V V E R L S  
5 ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250  
D A E R H G H T V L A L V R G S A  
GCTAACTCCGACGGCGGTGCAACGGTCTGTGCGGCGCGAACGGCCCCCTC 1300  
A N S D G A S N G L S A P N G P S  
CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350  
10 Q E R V I H Q A L A N A K L T P  
CCGATGTTCGACGCGGTTCGAGGCGCACGGCACCGGCACCGCCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D R A T  
15 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500  
P L L L G S L K S N I G H A Q A  
CGTCAGGGGTTCGCCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCGCACGTGCGACTG 1600  
20 E I P P T L H A D E P S P H V D W  
GACGGCCGGTGCCTGCGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGGA 1650  
T A G A V E L L T S A R P W P G  
CCGGTTCGCCCGCGCGCTGCCGTCTCGTCTGCGGCGTGAGCGGCACG 1700  
T G R A R A V S S F G V S G T  
25 AACGCCCACATCATCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750  
N A H I I L E A G P V K T G P V E  
GGCAGGAGCGATCGAGGCAGGACCGGTGCAAGTAGGACCGGTGAGGCTG 1800  
A G A I E A G P V E V G P V E A  
GACCGCTCCCCGCGGCGCGCGCTCAGCACCGGGCGAAGACCTTCCGCTG 1850  
30 G P L P A A P P S A P G E D L P L  
CTCGTGTTCGGCGCGTTCCCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
L V S A R S P E A L D E Q I G R L  
GCGCGCCTATCTCGACACCGGCCCGGGCGTTCGACCGGGCGGCGTGGCGC 1950  
R A Y L D T G P G V D R A A V A  
35 AGACACTGGCCCGCGGTACGCACTTACCCACCGGGCCGTACTGCTCGGG 2000  
Q T L A R R T H F T H R A V L L G  
GACACCGTTCATCGGCGCTCCCCCGCGGACCGGCGGACGAACTCGTCTT 2050  
D T V I G A P P A D Q A D E L V F  
CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100  
40 V Y S G Q G T Q H P A M G E Q L  
CCGCCGCGTTCCCCGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTG 2150  
A A A F P V F A R I H Q Q V W D L  
CTCGATGTGCCGATCTGGAGGTGAACGAGACCGGTTACGCCAGCCGGC 2200  
L D V P E V N E T G Y A Q P A  
45 CCTGTTTCGCAATGCAGGTGGCTCTGTTTCGGGCTGCTGGAATCGTGGGGTG 2250  
L F A M Q V A L F G L L E S W G  
TACGACCGGACGCGGTGATCGGCCATTCCGTGGGTGAGCTTTCGGGCTGCG 2300  
V R P D A V I G H S V G E L A A A  
TATGTGTCCGGGGTGTGGTTCGTTGGAGGATGCCTGCACTTTGGTGTTCGGC 2350  
50 Y V S G V W S L E D A C T L V S A  
GCGGGCTCGTCTGATGCAGGCTCTGCCCGCGGGTGGGGTGTGCTGCTG 2400  
R A R L M Q A L P A G G V M V A  
TCCCGGTCTCGGAGGATGAGGCCCGGGCGGTGCTGGGTGAGGGTGTGGAG 2450  
V P V S E D E A R A V L G E G V E  
55 ATCGCCGCGGTCAACGGCCCGTCTGTCGGTGGTTCTCTCCGGTGTGAGGC 2500  
I A A V N G P S S V V L S G D E A  
CGCCGTGCTGCAGGCCGCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGA 2550  
A V L Q A A E G L G K W T R L A  
CCAGCCACGCGTTCCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC 2600  
60 T S H A F H S A R M E P M L E E F  
CGGGCGGTTCGCCAAGGCCTGACCTACCGGACCGCGCAGGTCTCCATGGC 2650  
R A V A E G L T Y R T P Q V S M A  
CGTTGGTGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGG 2700  
V G D Q V T T A E Y W V R Q V R

ACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTC 2750  
D T V R F G E Q V A S Y E D A V F  
GTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCCTGGTTCGACGGGTGTTCGC 2800  
V E L G A D R S L A R L V D G V A  
5 GATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGCGCCCTGGCCC 2850  
M L H G D H E I Q A A I G A L A  
ACCTGTATGTCAACGGCGTCAACGGTCACTGGCCCGCGCTCCTGGGCGAT 2900  
H L Y V N G V T V D W P A L L G D  
10 GTCGGGCAACACGGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCA 2950  
A P A T R V L D L P T Y A F Q H Q  
GCGCTACTGGCTCGAGTCGGCTCCCCCGGCCACGGCCGACTCGGGCCACC 3000  
R Y W L E S A P P A T A D S G H  
CCGTCTCGGCACCGGAGTCGCCGTCGCCGGGTGCCGGGCGGGTGTTC 3050  
P V L G T G V A V A G S P G R V F  
15 ACGGGTCCCGTGCCCCCGGTGCGGACCGCGCGGTGTTTCATCGCCGAAC 3100  
T G P V P A G A D R A V F I A E L  
GGCGCTCGCCGCGCCGACGCCACCGACTGCGCCACGGTCAACAGCTCG 3150  
A L A A D A A T D C A T V E Q L  
ACGTCACTCCGTGCCCGCGGATCCGCCCGCGGCAGGGCCACCGCGCAG 3200  
20 D V T S V P G G S A R G R A T A Q  
ACCTGGGTGCGATGAACCCGCGCCGACGGGCGGCGCGCTTACCGTCEA 3250  
T W V D E P A A D G R R R F T V H  
CACCCGCGTTCGGCGACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCC 3300  
T R V G D A P W T L H A E G V L  
25 GCCCCGGCGCGTGCCCCAGCCCCGAAGCGTTCGACACCGCCTGGCCCCCG 3350  
R P G R V P Q P E A V D T A W P P  
CCGGGCGCGGTGCCCGCGGACGGGCTGCCCGGGCGTGGCGACGCGCGGA 3400  
P G A V P A D G L P G A W R R A D  
CCAGGTCTTCGTGCAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCAC 3450  
30 Q V F V E A E V D S P D G F V A  
ACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGGCGACGGGAGCCGC 3500  
H P D L D A V F S A V G D G S R  
CAGCCGACCGGATGGCGCGACCTCGCGGTGCACGCGTCGGACGCCACCGT 3550  
Q P T G W R D L A V H A S D A T V  
35 GCTGCGCGCCTTCCTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCG 3600  
L R A C L T R R D S G V V E L A  
CCTTCGACGGTGCCGGAATGCCGGTGTCAACCGCGGAGTCGGTGACGCTG 3650  
A F D G A G M P V L T A E S V T L  
GGCGAGGTGCGGTGCGGCGGATCCGACGAGTCGGACGGTCTGCTTCG 3700  
40 G E V A S A G G S D E S D G L L R  
GCTTGAGTGGTTGCCGGTGGCGGAGGCCACTACGACGGTGCCGACGAGC 3750  
L E W L P V A E A H Y D G A D E  
TGCCCCGAGGGCTACACCCTCATCACCGCCACACACCCCGACGACCCCGAC 3800  
L P E G Y T L I T A T H P D D P D  
45 GACCCACCAACCCCAACAACACCCACCGACCCACACACAAACCAC 3850  
D P T N P H N T P T R T H T Q T T  
ACGCGTCTTCACCGCCCTCCAACACCACCTCATCACCAACCAACCAACCC 3900  
R V L T A L Q H H L I T T N H T  
TCATCGTCCACACCACCGACCCCCAGGCGCGCGCTCACCGGCCTC 3950  
50 L I V H T T T D P P G A A V T G L  
ACCCGACCGCACAAAACGAACACCCCGCGCATCCACCTCATCGAAAC 4000  
T R T A Q N E H P G R I H L I E T  
CCACCACCCCAACCCCACTCCCCCTACCCAACTCACCACTCCACC 4050  
H H P H T P L P L T Q L T T L H  
55 AACCCACCTACGCCTACCAACAACACCTCCACACCCCCACCTCACC 4100  
Q P H L R L T N N T L H T P H L T  
CCCATCACCAACCAACAACCAACCAACCAACCAACCAACCAACCAACCAAC 4150  
P I T T H H N T T T T T P N T P P  
CCTCAACCCCAACCAACCACTCATCACCGGCGGCTCCGGCACCTCG 4200  
60 L N P N H A I L I T G G S G T L  
CCGGCATCTCGCCCGCCACCTCAACCAACCCCAACCACTACCTCCTCTCC 4250  
A G I L A R H L N H P H T Y L L S  
CGCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 4300  
R T P P P P T T P G T H I P C D L

CACCGACCCCAACCAATCACCCAAGCCCTCACCCACATACCACAACCCC 4350  
 T D P T Q I T Q A L T H I P Q P  
 TCACCGGCATCTTCCACACCGCCGCCACCCTCGACGACGCCACCCTCACC 4400  
 L T G I F H T A A T L D D A T L T  
 5 AACCTCACCCCAACACCTCACCACCACCCTCCAACCCAAAGCCGACGC 4450  
 N L T P Q H L T T T L Q P K A D A  
 CGCCTGGCACCTCCACCACCACACCCAAAACCAACCCCTCACCCACTTCG 4500  
 A W H L H H T Q N Q P L T H F  
 TCCTCTACTCCAGCGCCGCCACCCTCGGCAGCCCCGGCCAAGCCAAC 4550  
 10 V L Y S S A A A T L G S P G Q A N  
 TACGCCGCCGCCAACGCCTTCTCGACGCCCTCGCCACCCACCGCCACAC 4600  
 Y A A A N A F L D A L A T H R H T  
 CCAAGGACAACCCGCCACCACCATCGCCTGGGGCATGTGGCACACCACCA 4650  
 Q G Q P A T T I A W G M W H T T  
 15 CCACACTCACCAGCCAACTCACCAGACGCGACCGCGACCGCATCCGCCGC 4700  
 T T L T S Q L T D S D R D R I R R  
 GGCGGCTTCTCGCGATCTCGGACGACGAGGGCATGC  
 G G F L P I S D D E G M

20 The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of  
 module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 M R L Y E A A R R T G S P V V V  
 GCGGCCGCGCTCGACGACGCGCCGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 25 A A L D A P D V P L L R G L R  
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
 R T T V R R A A V R E R S L A D  
 GCTCGCCGTGCTGCCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTCTG 200  
 R S P C C P T T S A P T P P S R S  
 30 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 CCCGGCGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300  
 P A T T T F K E L G I D S L T A  
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
 35 V Q L R N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400  
 T A V F D F P T P R A L A A R L G  
 CGAGAGCTGGCCGGTACCCGCGCGCCGTCGCGGCCCGGACCGCGGCCA 450  
 D E L A G T R A P V A A R T A A  
 40 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 L P G G V A S P Q E L W R L V A S  
 CGGCACCGACGCCATCACGGAGTTCCCGCGGACCGCGGCTGGGACGTGG 600  
 45 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCGGACCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 D A L Y D P D P D A I G K T F V R  
 CACGGCGGCTTCTCGACGGTGGCACC GGCTTCGACGCGGCGTTCTTCGG 700  
 H G G F L D G A T G F D A A F F G  
 50 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q R V L  
 TGGAGACGTCCTGGGAGGCGTTTCGAAAGCGCGGGCATCACCCCGGACGCG 800  
 L E T S W E A F E S A G I T P D A  
 GCGCGGGGCGAGCAGACACCGGCGTGTTCATCGGCGCGTTCTCTACGGGTA 850  
 55 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGGACAGGGTTCGACAGACCA 900  
 G T G A D T N G F G A T G S Q T  
 GCGTGCTCTCCGGCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 S V L S G R L S Y F Y G L E G P S  
 60 GTCACGGTCGACACCGCCTGCTCGTCGTCGCTGCTGGTGGCCCTGCACCAGGC 1000  
 V T V D T A C S S S L V A L H Q A  
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTGGCGGGT 1050

G Q S L R S G E C S L A L V G G  
TCACGGTGATGGCGTCGCCCGGCGGATTCTCGTCGAGTTCTCCCGGCAGCGC 1100  
V T V M A S P G G F V E F S R Q R  
5 GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150  
G L A F D G R A K A F G A G A D G  
TACGAGCTTCGCCGAGGGCGCCGGTGCCTTGGTGGTTCGAGCGGCTCTCCG 1200  
T S F A E G A G A L V V E R L S  
ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250  
D A E R H G H T V L A L V R G S A  
10 GCTAACTCCGACGGCGCGTCAACGGTCTGTGGCGCCGAACGGCCCTC 1300  
A N S D G A S N G L S A P N G P S  
CCAGGAACGCGTCATCCACAGGCCCTCGCGAACGCGAACTCACCCCG 1350  
Q E R V I H Q A L A N A K L T P  
CCGATGTGACGCGGTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400  
15 A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D R A T  
GCCCCGTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500  
P L L L G S L K S N I G H A Q A  
20 CGTCAGGGGTGCGCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCCGACGTCGACTG 1600  
E L P P T L H A D E P S P H V D W  
GACGGCCGGTGCCGTGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGA 1650  
25 T A G A V E L L T S A R P W P G  
CCGGTCGCCCGCGCCGCGTCCGCTCTCGTCGTTTCGGCGTGAGCGGCACG 1700  
T G R P R R A A V S S F G V S G T  
AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAACGGGACCGGTCTGA 1750  
N A H I I L E A G P V K T G P V E  
30 GGCAGGAGCGATCGAGGCAGGACCGGTCTGAAGTAGGACCGGTCTGAGGCTG 1800  
A G A I E A G P V E V G P V E A  
GACCGCTCCCCGCGCGCCGCGTCCAGCACCGGGCGAAGACCTTCCGCTG 1850  
G P L P A P P S A P G E D L P L  
CTCGTGTGCGCGGTTCCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
35 L V S A R S P E A L D E Q I G R L  
GCGCGCTATCTCGACACCGGCCCGGGCGTCCAGCGGGCGGCCGTGGCGC 1950  
R A Y L D T G P G V D R A A V A  
AGACACTGGCCCGGCGTACGCACTTACCCACCGGCCGTACTGCTCGGG 2000  
Q T L A R R T H F T H R A V L L G  
40 GACACCGTCTCGGCGCTCCCCCGGACCGGCGAGGCGAAGTCTGCTCTT 2050  
D T V I G A P P A D Q A D E L V F  
CGTCCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100  
V Y S G Q G T Q H P A M G E Q L  
CCGATTCGTGCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGGC 2150  
45 A D S S V V F A E R M A E C A A A  
TTGCGCGAGTTCGTGGACTGGGATCTGTTACGGTCTTGATGATCCGGC 2200  
L R E F V D W D L F T V L D D P A  
GGTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCTGGGCGATGATGG 2250  
V V D R V D V V Q P A S W A M M  
50 TTTCCCTGGCCGCGGTGTGGCAGGCGGCCGTGTGCGGCCGGATGCGGTG 2300  
V S L A A V W Q A A G V R P D A V  
ATCGGCCATTTCGAGGGTGAGATCGCCGAGCTTGTGTGGCGGGTGGCGT 2350  
I G H S Q G E I A A A C V A G A V  
GTCACTACGCGATGCCGCCGCGATCGTGACCTTGGCGAGCCAGGCGATCG 2400  
55 S L R D A A R I V T L R S Q A I  
CCCCGGGCTTGGCGGGCGGGCGCGATGGCATCCGTGCCCCTGCCCGCG 2450  
A R G L A G R G A M A S V A L P A  
CAGGATGTCGAGCTGGTCGACGGGGCCTGGATCGCCGCCCAACGGGCC 2500  
Q D V E L V D G A W I A A H N G P  
60 CGCCTCCACCGTGATCGCGGCGACCCCGAAGCGGTGACCATGTCTCTCA 2550  
A S T V I A G T P E A V D H V L  
CCGCTCATGAGGCACAAGGGGTGCGGGTGGCGGCGATCACCGTCGACTAT 2600  
T A H E A Q G V R V R R I T V D Y  
GCCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAACTACTCGACAT 2650

A S H T P H V E L I R D E L L D I  
 CACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTCGACCG 2700  
 T S D S S S Q T P L V P W L S T  
 TGGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTACTGGTACCGG 2750  
 5 V D G T W V D S P L D G E Y W Y R  
 AACCTGCGTGAACCGGTGCGTTTCCACCCCGCCGTCAGCCAGTTGCAGGC 2800  
 N L R E P V G F H P A V S Q L Q A  
 CCAGGGCGACACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGC 2850  
 Q G D T V F V E V S A S P V L L  
 10 AGGCGATGGACGACGATGTCGTACGGTTGCCACGCTGCGTCTGACGAC 2900  
 Q A M D D V V T V A T L R R D D  
 GGCGACGCCACCCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGG 2950  
 G D A T R M L T A L A Q A Y V H G  
 CGTCACCGTCTGACTGGCCCGCCATCCTCGGCACCACCACAACCCGGGTAC 3000  
 15 V T V D W P A I L G T T T T R V  
 TGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCG 3050  
 L D L P L T Y A F Q H Q R Y W L E S  
 GCTCCCCCGGCACGCGGACTCGGGCCACCCCGTCTCGGCACCCGGAGT 3100  
 A P P A T A D S G H P V L G T G V  
 20 CGCCGTCGCCGGGTGCGCCGGCGGGTGTTCACGGGTCCCGTGCCCGCCG 3150  
 A V A G S P G R V F T G P V P A  
 GTGCGGACCGCGCGGTGTTTCATCGCCGAACTGGCGCTCGCCGCCGCCGAC 3200  
 G A D R A V F I A E L A L A A A D  
 GCCACCGACTGCGCCACGGTGAACAGCTCGACGTACCTCCGTGCCCCG 3250  
 25 A T D C A T V E Q L D V T S V P G  
 CGGATCCGCCCGCGGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCG 3300  
 G S A R G R A T A Q T W V D E P  
 CCGCCGACGGGCGGCGCGCTTACCGTCCACACCCGCGTGGCGACGCC 3350  
 A A D G R R R F T V H T R V G D A  
 30 CCGTGGACGCTGCACGCCGAGGGGGTCTCCGCCCGCGCGGTGCCCCA 3400  
 P W T L H A E G V L R P G R V P Q  
 GCCCGAAGCCGTGACACCCGCTGGCCCCCGCGGGCGCGGTGCCCGCGG 3450  
 P E A V D T A W P P P G A V P A  
 ACGGGCTGCCCCGGGCGTGCGACGCGCGGACCAGGTCTTCGTGAAGCC 3500  
 35 D G L P G A W R R A D Q V F V E A  
 GAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCGACCTGCTCGACGC 3550  
 E V D S P D G F V A H P D L L D A  
 GGTCTTCTCCGCGGTGCGGACGGGAGCCGCGACCGGATGGCGCG 3600  
 V F S A V G D G S R Q P T G W R  
 40 ACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCTGCCTCACC 3650  
 D L A V H A S D A T V L R A C L T  
 CGCCGCGACAGTGGTGTCTGAGGCTCGCCGCTTCGACGGTGCCGGAAT 3700  
 R R D S G V V E L A A F D G A G M  
 GCCGTGTCTACCGCGGAGTGGGTGACGCTGGGCGAGGTGCGGTGCGCAG 3750  
 45 P V L T A E S V T L G E V A S A  
 GCGGATCCGACGAGTGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTG 3800  
 G G S D E S D G L L R L E W L P V  
 GCGGAGGCCCCACTACGACGGTGCCGACGAGTGGCCGAGGGCTACACCCT 3850  
 A E A H Y D G A D E L P E G Y T L  
 50 CATACCGCCACACACCCCGACGACCCCGACGACCCCAACCCCA 3900  
 I T A T H P D D P D D P T N P H  
 ACACACCCACACGCACCCACACACAAACACACGCTCCTACCGCCCTC 3950  
 N T P T R T H T Q T T R V L T A L  
 CAACACCACCTCATCACCACCAACCACACCTCATCGTCCACACCACCAC 4000  
 55 Q H H L I T T N H T L I V H T T T  
 CGACCCCCAG3CGCCGCGGTACCGGCCTACCCGACCCGACAAAACG 4050  
 D P P G A A V T G L T R T A Q N  
 AACACCCGCGCATCCACCTCATCGAAACCCACACCCCAACCCCA 4100  
 E H P G R I H L I E T H H P H T P  
 60 CTCCCCCTACCCAACTCACCACCTCCACCAACCCCACTACGCCTCAC 4150  
 L P L T Q L T T L H Q P H L R L T  
 CAACAACACCTCCACACCCCCACCTCACCCTCATCACCACCCACCACA 4200  
 N N T L H T P H L T P I T T H H  
 ACACCACCAACCAACCCCAACACCCCAACCCCTCAACCCCAACCAACGCC 4250

```

N T T T T T P N T P P L N P N H A
ATCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGCCCGCCA 4300
I L I T G G S G T L A G I L A R H
CCTCAACCACCCCCACACCTACCTCCTCTCCCGCACACCACCACCCCCA 4350
5   L N H P H T Y L L S R T P P P P
CCACACCCGGCACCCACATCCCCTGCGACCTCACCAGCCCCACCCAAATC 4400
T T P G T H I P C D L T D P T Q I
ACCCAAGCCCTCACCCACATAACCACAACCCCTCACCGGCATCTTCCACAC 4450
T Q A L T H I P Q P L T G I F H T
10  CGCCGCCACCCTCGACGACGCCACCCTCACCAACCTCACCCCCAACACC 4500
A A T L D D A T L T N L T P Q H
TCACCACCACCCTCCAACCCAAAGCCGACGCCGCCTGGCACCTCCACCAC 4550
L T T L Q P K A D A A W H L H H
CACACCCAAAACCAACCCCTCACCCACTTCGTCTCTACTCCAGCGCCGC 4600
15  H T Q N Q P L T H F V L Y S S A A
CGCCACCCTCGGCAGCCCCGGCCAAGCCAACCTACGCCGCCGCCAACGCCT 4650
A T L G S P G Q A N Y A A A N A
TCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACC 4700
F L D A L A T H R H T Q G Q P A T
20  ACCATCGCTGGGGCATGTGGCACACCACCACACTCACCAGCCAACT 4750
T I A W G M W H T T T T L T S Q L
CACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCTGCCGATCT 4800
T D S D R D R I R R G G F L P I
CGGACGACGAGGGCATGC
25  S D D E G M

```

### Example 3

#### Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from *Streptomyces* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding sequences have been replaced by either the *rapAT3* (the AT domain from module 3 of the rapamycin PKS), *rapAT12*, *eryAT1* (the AT domain from module 1 of the erythromycin (DEBS) PKS), or *eryAT2* coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the *rapAT12* replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites *SacI* and *SphI* (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique *SacI* and *SphI* restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique *Bgl* II and *Nsi* I sites by ligation to synthetic linkers (described in

the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8 sequences were then amplified using primers, described above, that introduced either an *AvrII* site or an *NheI* site at two different KS/AT boundaries and an *XhoI* site at the AT/DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the *BamHI* and *PstI* sites of the KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Heterologous AT	Enzyme	Location of Engineered Site
FK-506 AT8 (hydroxymalonyl)	<i>AvrII</i>	GGCCGT <u>ccgcgc</u> CGTGCGGCGGTCTCGTCGTTTC G R P R R A A V S S F
	<i>NheI</i>	ACCCAGCATCCCGCGATGGGTGAGCG <u>gctcgc</u> C T Q H P A M G E R L A
	<i>XhoI</i>	TACGCCTTCCAGCGGCGGCCCTACTGG <u>atcgag</u> Y A F Q R R P Y W I E
rapamycin AT3 (methylmalonyl)	<i>AvrII</i>	GACCGG <u>ccccgc</u> CGGGCGGGCGTGTCTCGTCCTTC D R P R R A G V S S F
	<i>NheI</i>	TGGCAGTGGCTGGGGATGGGCAGTGC <u>cctgcgc</u> G W Q W L G M G S A L R
	<i>XhoI</i>	TACGCCTTCCAACACCAGCGGTACTGG <u>gtcgag</u> Y A F Q H Q R Y W V E
rapamycin AT12 (malonyl)	<i>AvrII</i>	GGCCGAG <u>gcgcgc</u> CGGGCAGGCGTGTCTCGTCCTTC G R A R R A G V S S F
	<i>NheI</i>	TCGCAGCGTGTGGCATGGGTGAGGA <u>actggc</u> C S Q R A G M G E E L A
	<i>XhoI</i>	TACGCCTTCCAGCACCAGCGCTACTGG <u>gtcgag</u> Y A F Q H Q R Y W L E
DEBS AT1 (methylmalonyl)	<i>AvrII</i>	GCGCGA <u>accgcgc</u> CGGGCGGGGTCTCGTCGTTTC A R P R R A G V S S F
	<i>NheI</i>	TGGCAGTGGGCGGGCATGGCCGTCGA <u>cctgct</u> C W Q W A G M A V D L L
	<i>XhoI</i>	TACCCGTTCCAGCGCGAGCGCGTCTGG <u>gtcgaa</u> Y P F Q R E R V W L E
DEBS AT2 (methylmalonyl)	<i>AvrII</i>	GACGGG <u>gtgcgc</u> CGGGCAGGTGTGTCTGGCGTTTC D G V R R A G V S A F GCCCAGTGGGAAGGCATGGCGCGGA <u>attatt</u> G



	<i>NheI</i>	A Q W E G M A R E L L
		TATCCTTTCCAGGGCAAGCGGTTCTGG <u>ctgctg</u>
	<i>XhoI</i>	Y P F Q G K R F W L L

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

5 CCGGCGCCGTCGAACTGCTGACGTCGGCCCCGGCCGTGGCCCCGAGACCGACCGGccacggc  
 A G A V E L L T S A R P W P E T D R P R  
 GTGCCGCCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCACGTCATCCTGGAGGCCG  
 R A A V S S F G V S G T N A H V I L E A  
 GACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGACCTTCCCCTGCTGGTGTGG  
 10 G P V T E T P A A S P S G D L P L L V S  
 CACGCTCACCGGAAGCGCTCGACGAGCAGATCCGCCGACTGCGCGCCTACCTGGACACCA  
 A R S P E A L D E Q I R R L R A Y L D T  
 CCCCCGATCGACCGGGTGGCCGTGGCACAGACGCTGGCCCCGGCGCACACACTTCGCCC  
 T P D V D R V A V A Q T L A R R T H F A  
 ACCGCGCCGTGCTGCTCGGTGACACCGTCATCACCACACCCCCCGCGGACCGGCCCGACG  
 15 H R A V L L G D T V I T T P P A D R P D  
 AACTCGTCTTCGTCTACTCCGGCCAGGGCACCCAGCATCCCGCGATGGGCGAGCAgctcg  
 E L V F V Y S G Q G T Q H P A M G E Q L  
 CGCGCGCCATCCCGTGTTCGCCGACGCCTGGCATGAAGCGCTCCGCCGCCTTGACAACC  
 20 A A A P P V F A D A W H E A L R R L D N

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-520 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

25 TCCTCGGGGCTGGGTACGGCACGACGCGGATGTGCCCGCGTACGCGTTCCAACGGCGGC  
 I L G A G S R H D A D V P A Y A F Q R R  
 ACTACTGGatcgagTCGGCACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGTGGGCT  
 H Y W I E S A R P A A S D A G H P V L G

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-506 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

30 TCGGCCAGGCCGTGGCCGCGGACCGGCCGTccgcgcCGTGGCGCGGTCTCGTCTCGGG  
 S A R P W P R T G R P R R A A V S S F G  
 GTGAGCGGCACCAACGCCACATCATCCTGGAGGCCGACCCGACAGGAGGAGCCGTCG  
 35 V S G T N A H I I L E A G P D Q E E P S  
 GCAGAACCGGCCGGTGACCTCCCGTCTCGTGTGGCACGGTCCCGGAGGCACTGGAC  
 A E P A G D L P L L V S A R S P E A L D  
 GAGCAGATCGGGCGCCTGCGCGACTATCTCGACGCCGCCCCCGCGTGGACCTGGCGGCC  
 E Q I G R L R D Y L D A A P G V D L A A  
 40 GTGGCGCGGACACTGGCCACGCGTACGCACTTCTCCACCGCGCCGTACTGCTCGGTGAC  
 V A R T L A T R T H F S H R A V L L G D  
 ACCGTCATCACCGCTCCCCCGTGGAACAGCCGGGCGAGCTCGTCTTCGTCTACTCGGGA  
 T V I T A P P V E Q P G E L V F V Y S G  
 CAGGGCACCCAGCATCCCGCGATGGGTGAGCGgctcgCGCAGCCTTCCCCGTGTTCGCC  
 45 Q G T Q H P A M G E R L A A A F P V F A  
 GACCCGACGTACCCGCCTACGCCTTCCAGCGCGGCCCTACTGGATCGAGTCCGCGCCG  
 D P D V P A Y A F Q R R P Y W I E S A P

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-506 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

GACCCGACGTACCCGCCTACGCCTTCCAGCGCGGCCCTACTGGatcgagTCCGCGCCG  
 D P D V P A Y A F Q R R P Y W I E S A P

Example 4Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506 and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or methyl. These derivatives are produced in recombinant host cells of the invention that express recombinant PKS enzymes that produce the derivatives. These recombinant PKS enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the present invention provides recombinant PKS enzymes in which the AT domains of both modules 7 and 8 have been changed. The table below summarizes the various compounds provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
15	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound -- FK-506
20	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
25	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound -- FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
30	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and/or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module.

Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domains but also those in which one of the modules is converted to an ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

#### Example 6

##### Neurotrophic Compounds

The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation. These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of FK-506. The 18-hydroxy compounds of the invention

can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

5           The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45  $\mu$ L of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64  $\mu$ L) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with  
10   brine, and the organic phase dried over magnesium sulfate. Removal of solvent *in vacuo* and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53  $\mu$ L of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is  
15   cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is  
20   dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-18-hydroxyl derivatives of the compounds described in Examples 1 - 4.

25           Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai *et al.*, Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, *FEBS Letters* 316(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[*S*]-OH and C18-[*R*]-OH enantiomers, with  
30   the *R* enantiomer showing a somewhat lower IC<sub>50</sub>, which may be preferred in some applications. See Kawai *et al.*, *supra*. Another preferred protocol is described in Umbreit and Sharpless, 1977, JACS 99(16): 1526-28, although it may be preferable to use 30 equivalents each of SeO<sub>2</sub> and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of  
5 illustration and not limitation of the following claims.

Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthetase, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.
- 5
2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 10
3. The isolated nucleic acid of claim 1 that encodes an open reading frame, said open reading frame comprising coding sequences for two or more extender modules, each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 15
4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 20
5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.
- 25
6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.
- 30
7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid pKOS065-M21.
- 35
8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase, FK-506 polyketide synthase, or erythromycin polyketide synthase.

9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.

5

10. The method of claim 9, wherein said host cell is a *Streptomyces* host cell.

11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.

10

12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.

15

13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

20

14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

25

15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.

30

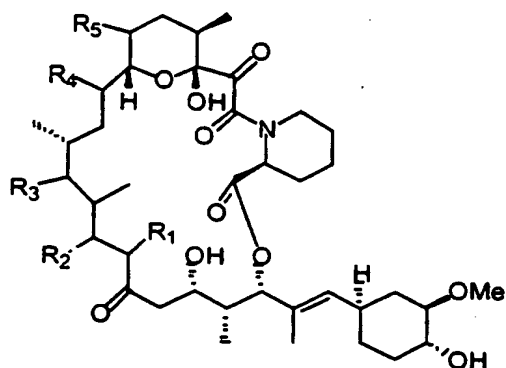
16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.

17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

35



18. A polyketide having the structure



- 5 wherein,  $R_1$  is hydrogen, methyl, ethyl, or allyl;  $R_2$  is hydrogen or hydroxyl, provided that when  $R_2$  is hydrogen, there is a double bond between C-20 and C-19;  $R_3$  is hydrogen or hydroxyl;  $R_4$  is methoxyl, hydrogen, methyl, or ethyl; and  $R_5$  is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.

10

19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.

20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.

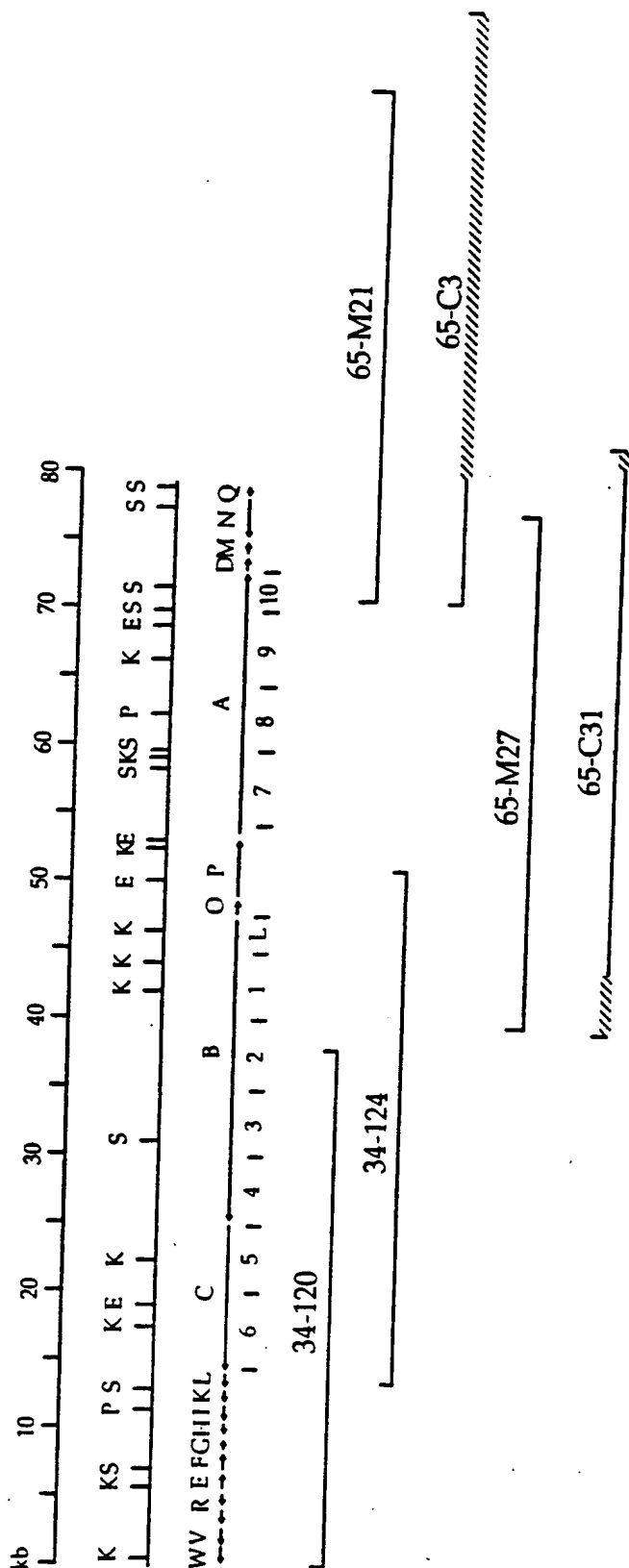


Figure 1

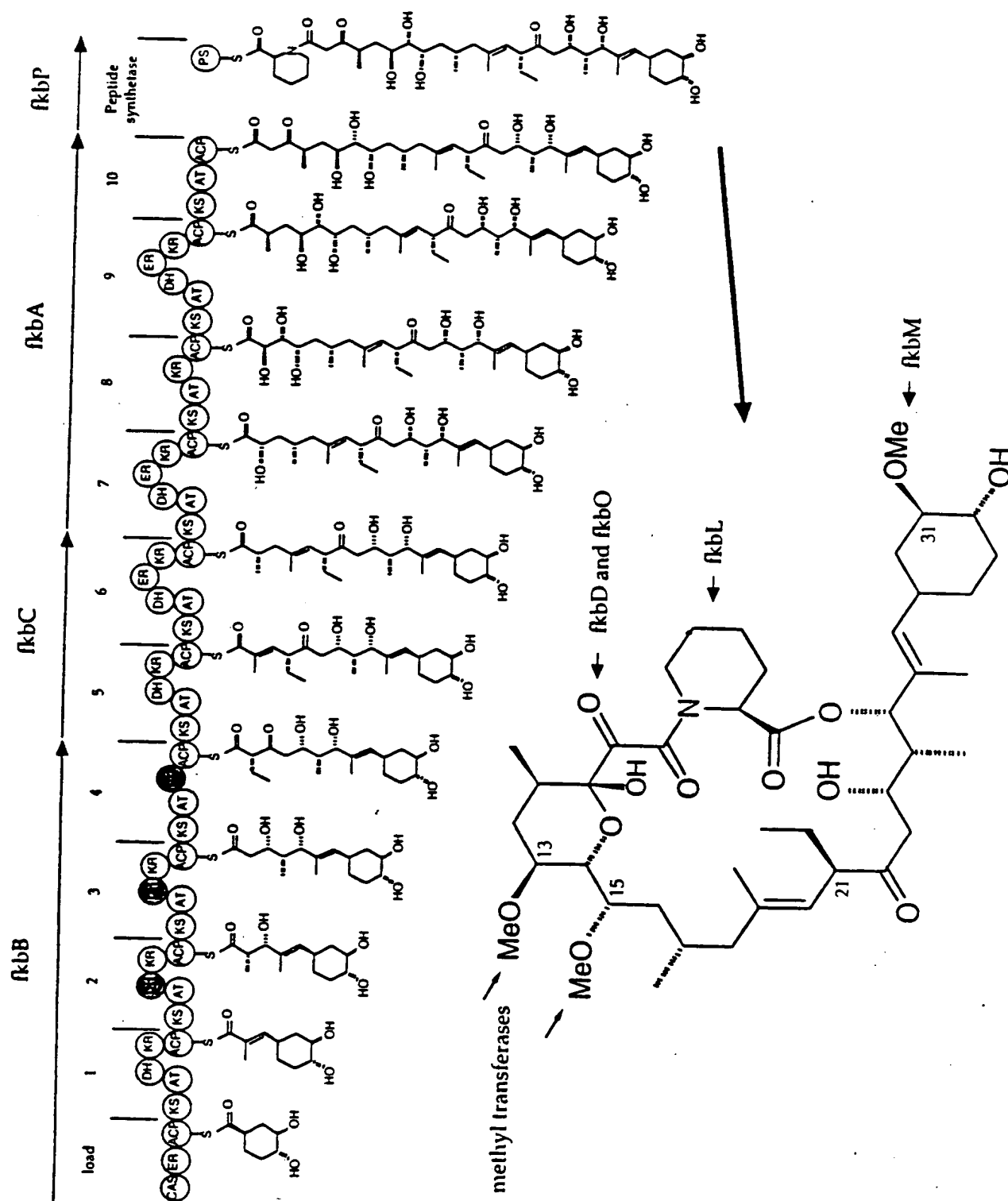


Figure 2

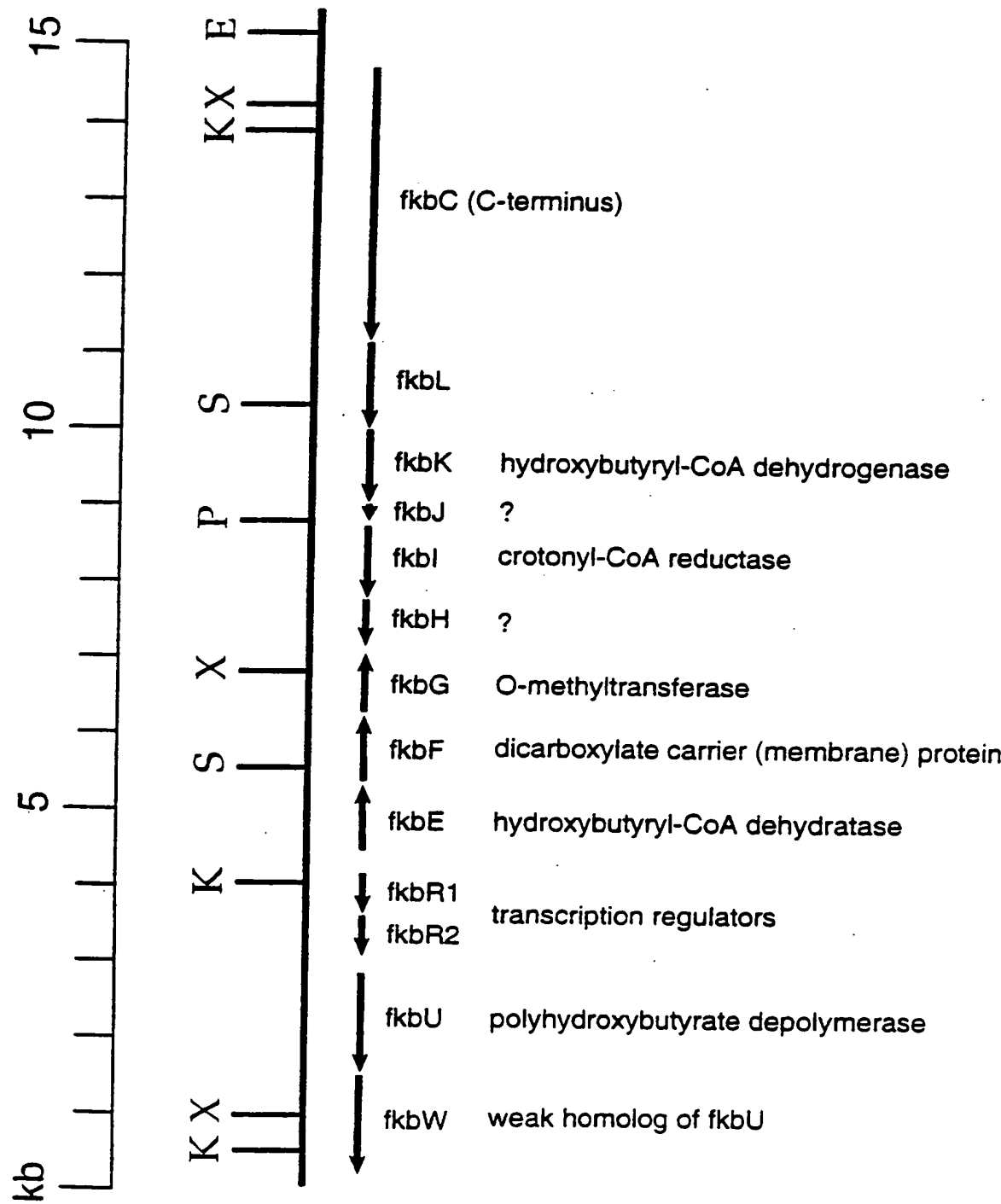


Figure 3

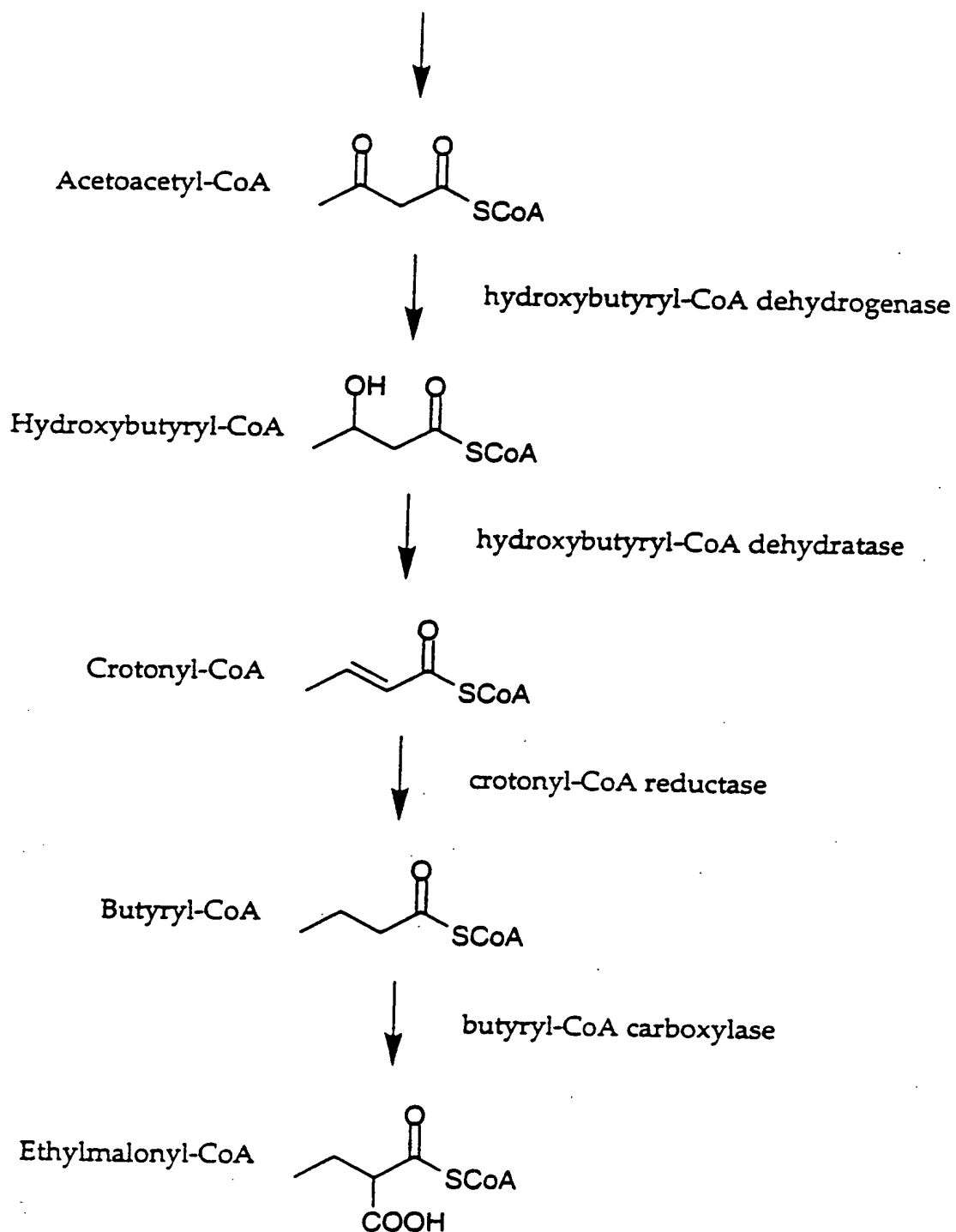


Figure 4

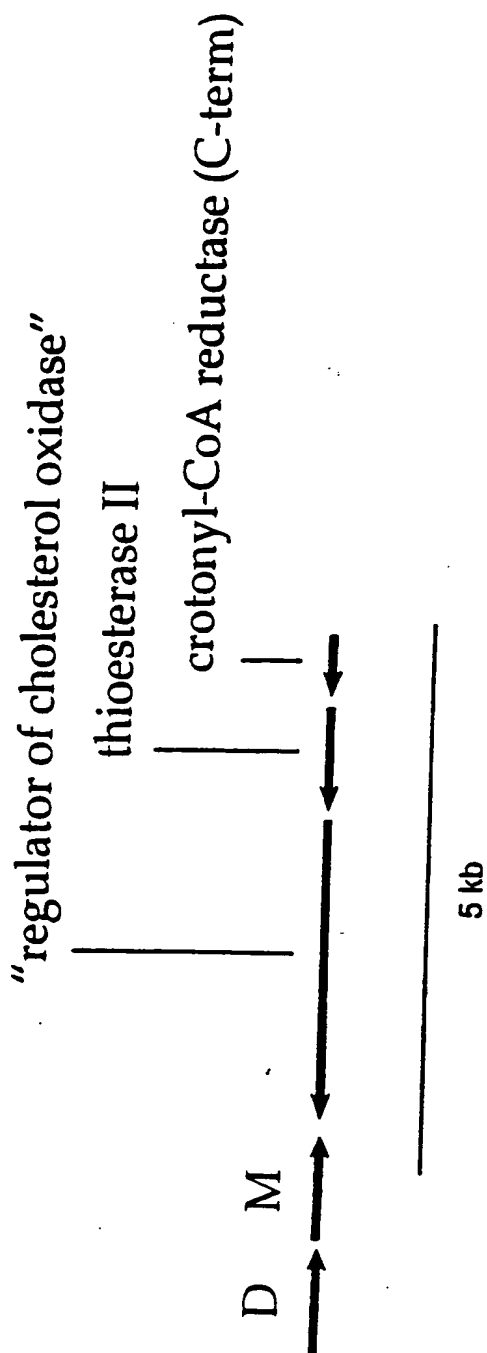


Figure 5

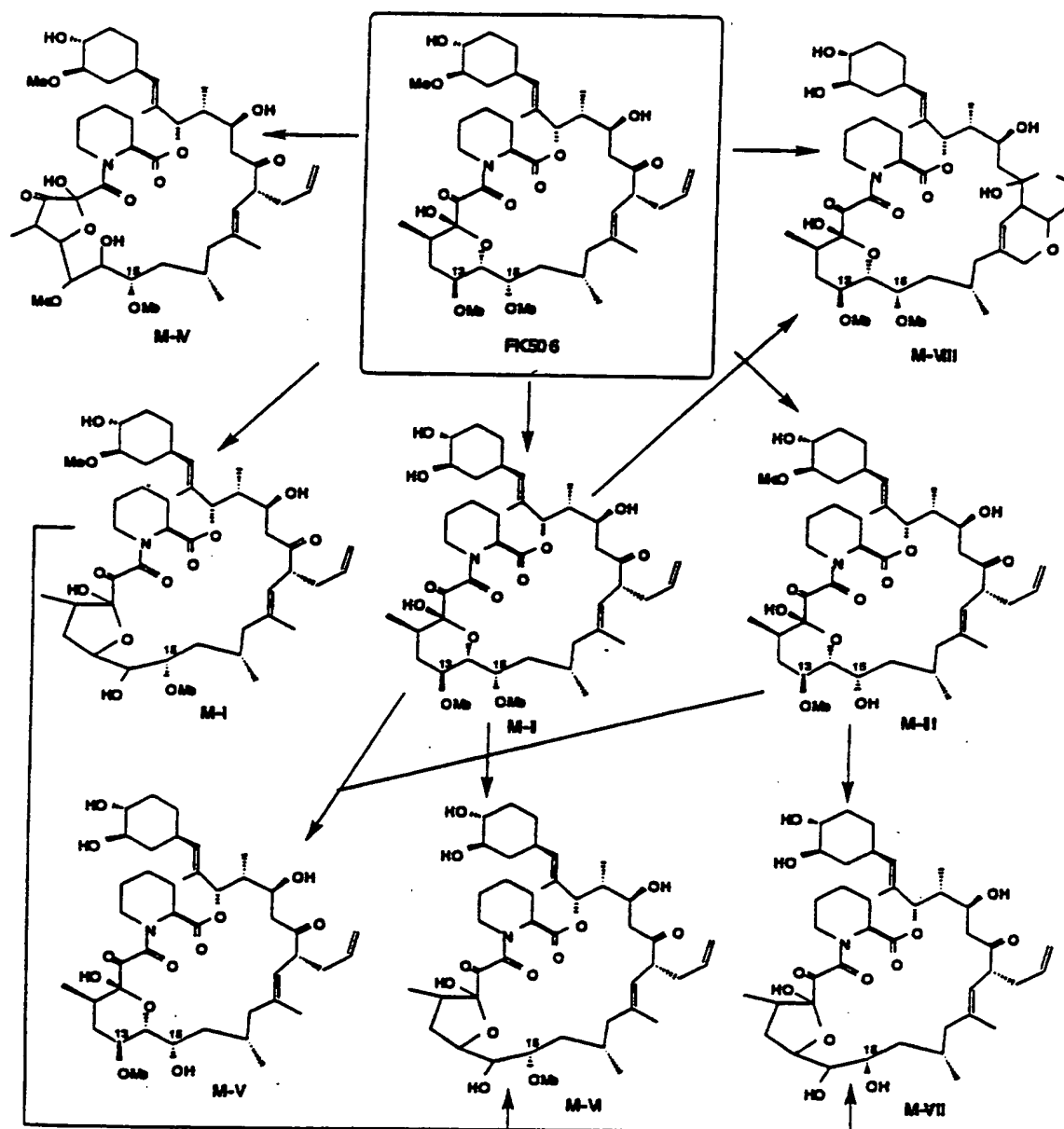


Figure 6

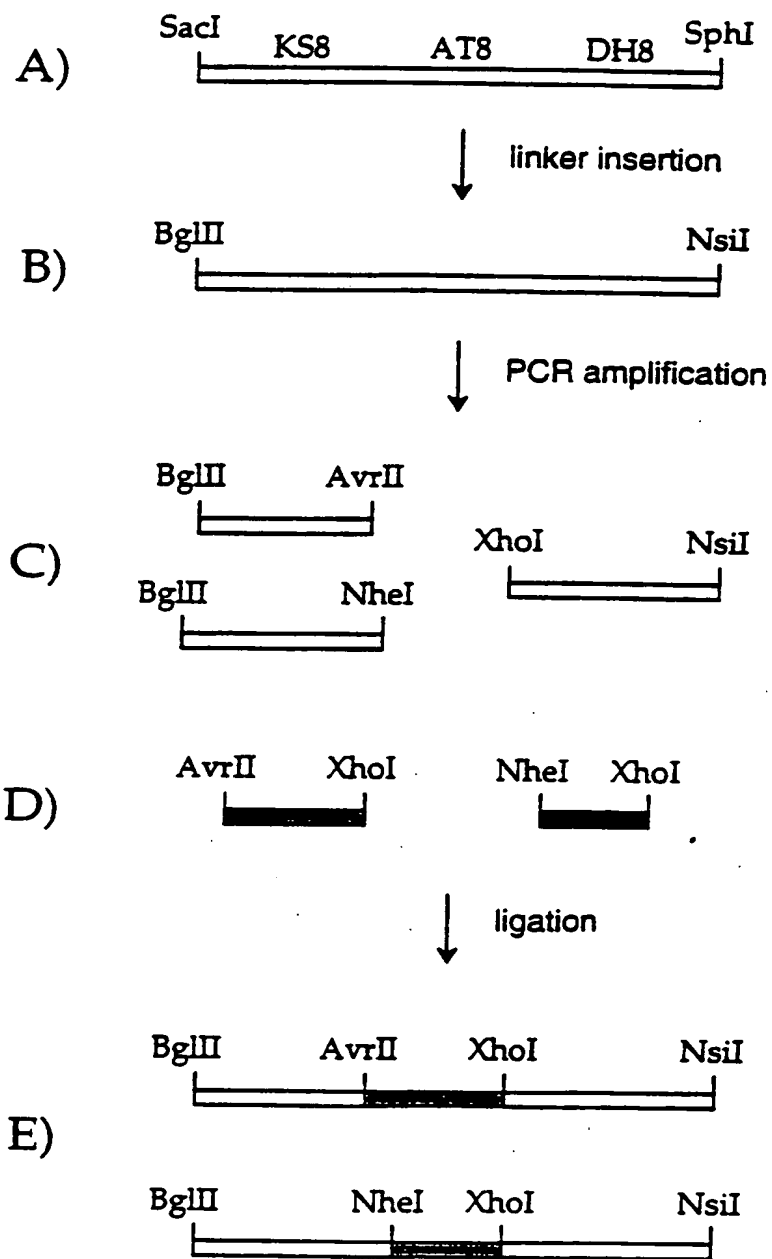
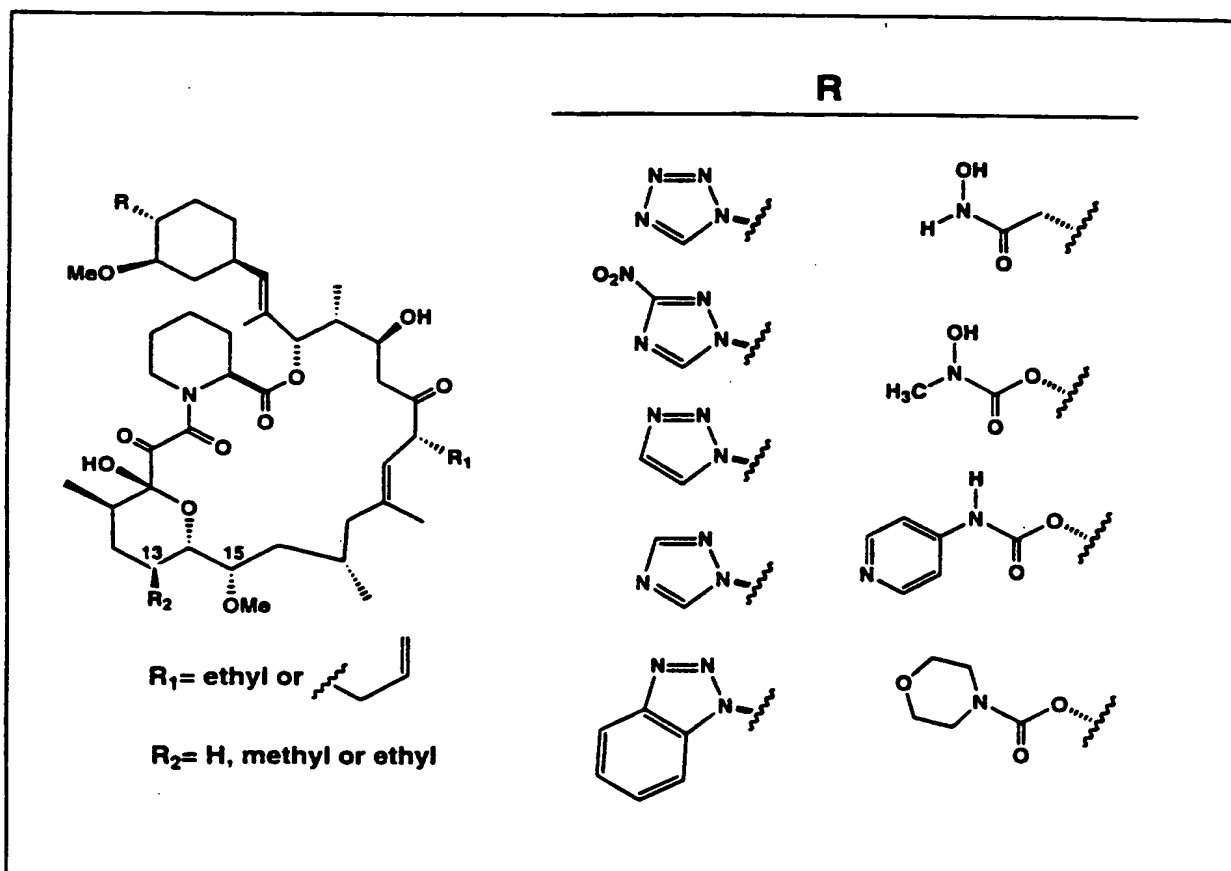
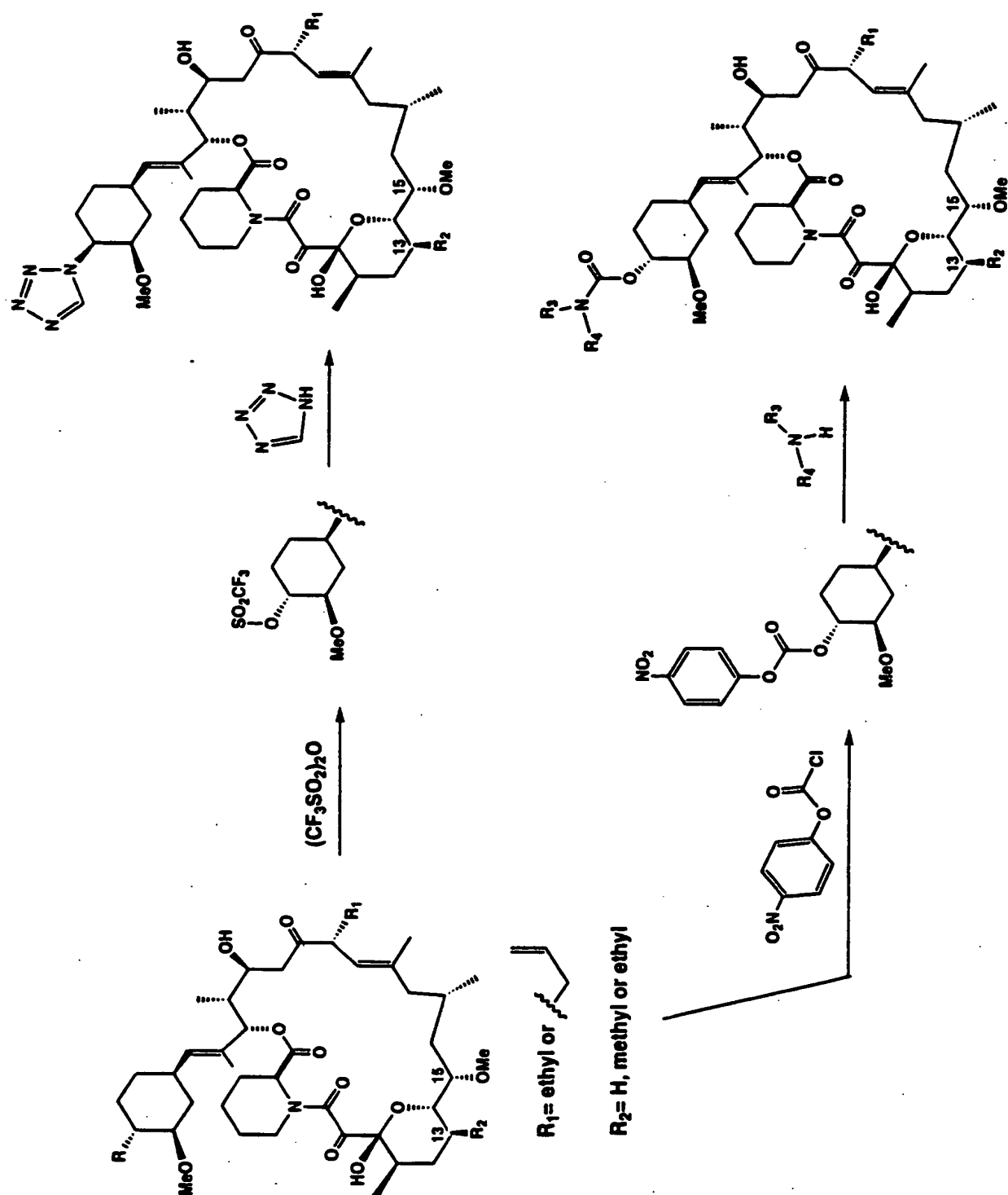


Figure 7





**Figure 8**  
**Part A**

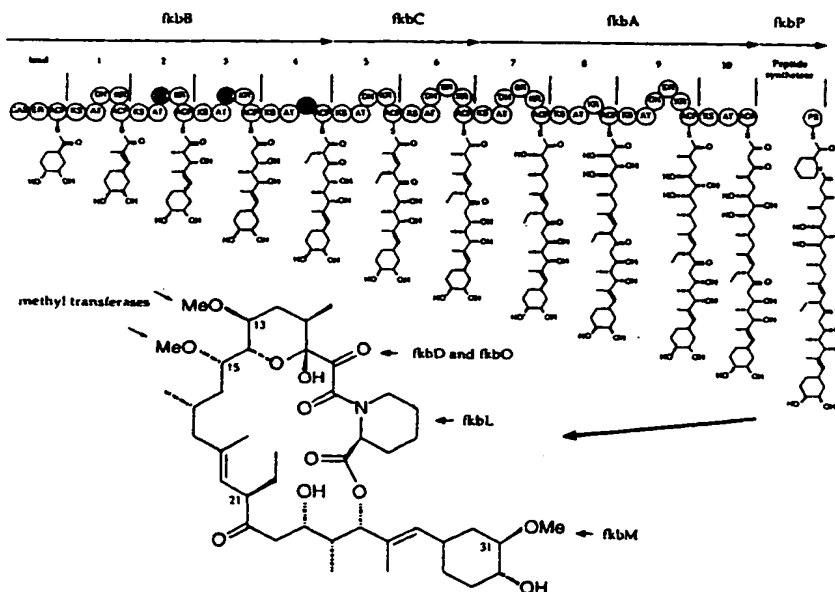
Figure 8  
Part B



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)</b>		A2	(11) International Publication Number: <b>WO 00/20601</b>
(21) International Application Number: <b>PCT/US99/22886</b>		(43) International Publication Date: <b>13 April 2000 (13.04.00)</b>	
(22) International Filing Date: <b>1 October 1999 (01.10.99)</b>		(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).	
(30) Priority Data: 60/102,748      2 October 1998 (02.10.98)      US 60/123,810      11 March 1999 (11.03.99)      US 60/139,650      17 June 1999 (17.06.99)      US		(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).		<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>	
(72) Inventors; and (75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).			

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



## (57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	JP	Japan	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	KE	Kenya	NL	Netherlands	VN	Viet Nam
CG	Congo	KG	Kyrgyzstan	NO	Norway	YU	Yugoslavia
CH	Switzerland	KP	Democratic People's Republic of Korea	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire			PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS  
THEREFOR

5

Field of the Invention

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fields of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

10

Background of the Invention

15

Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes.

Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline, erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

20

25

This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either to produce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5,063,155; 5,098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5,843,718; and Fu *et al.*, 1994, *Biochemistry* 33:

30

9321-9326; McDaniel *et al.*, 1993, *Science* 262: 1546-1550; and Rohr, 1995, *Angew. Chem. Int. Ed. Engl.* 34(8): 881-888, each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal peptide synthetase (NRPS). Two major types of PKS enzymes are known; these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as *eryAI*, *eryAII*, and *eryAIII*. See Caffrey *et al.*, 1992, *FEBS Letters* 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS<sup>Q</sup>, where the superscript letter is the abbreviation for the amino acid, glutamine, that is

present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior module, optionally performing one or more additional functions, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS, AT, ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom, producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta

keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS, AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, a DH domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activities, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes; these enzymes add carbohydrate groups or methyl groups, or make other modifications, i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of erythromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypeptides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those



taken from other sources. A genetically engineered PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence  
5 alignments also have revealed linker regions between the catalytic domains and at the N- and C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and between the individual polypeptides that comprise the PKS complex. One can thus view the  
10 linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT replacement, one can  
15 thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that known polyketides can be produced more  
20 effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes.  
25 The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present invention helps  
30 meet the need for such compounds as well.

#### Summary of the Invention

In one embodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3, pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the  
5 various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

The invention also provides isolated nucleic acids that encode all or part of one or  
10 more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors  
15 containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the  
20 FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS  
25 domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-506 or rapamycin.

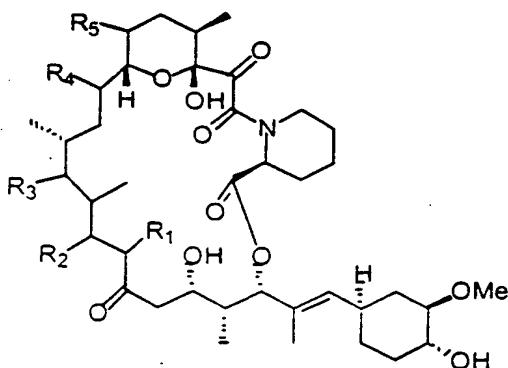
30 In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis.

The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

5 In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynthesis. The invention also provides recombinant  
10 nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to  
15 FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the  
20 invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppression activities.

25 Thus, the invention provides polyketides having the structure:



wherein, R<sub>1</sub> is hydrogen, methyl, ethyl, or allyl; R<sub>2</sub> is hydrogen or hydroxyl, provided that when R<sub>2</sub> is hydrogen, there is a double bond between C-20 and C-19; R<sub>3</sub> is hydrogen or hydroxyl; R<sub>4</sub> is methoxyl, hydrogen, methyl, or ethyl; and R<sub>5</sub> is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and their brief description below, together with the detailed description, examples, and claims that follow.

#### Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbc*. Immediately under the third line are numbered segments showing where the loading module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various cosmids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3, and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the

stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chain on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol.* 39:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fk bD*, *fk bM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fk bN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fk bQ* (a type II thioesterase, which can increase polyketide production levels), and *fk bS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

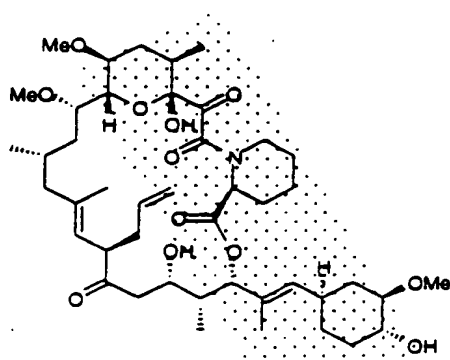
Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

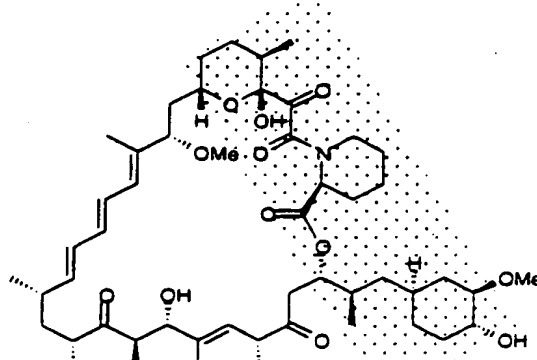
### Detailed Description of the Invention

5        Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see Holt *et al.*, 1993, *JACS* 115:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart, kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the unapproved use of tacrolimus for other conditions, including alopecia universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

20        The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.



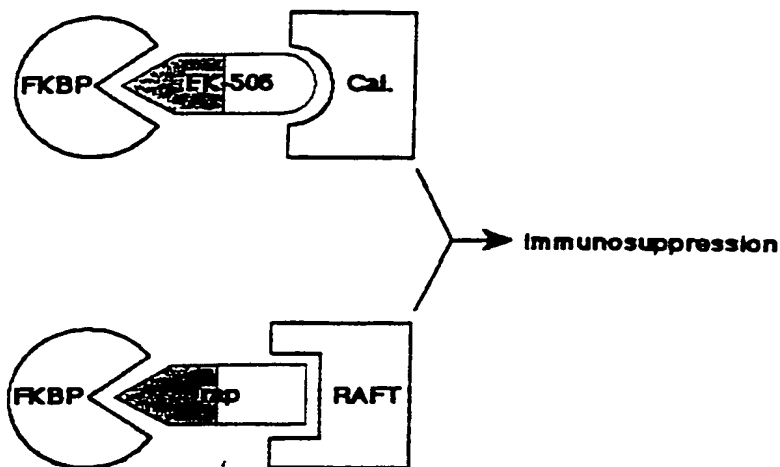
FK-506



Rapamycin

FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

These compounds act through initial formation of an intermediate complex with protein "immunophilins" known as FKBP (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules, known as the "FKBP-binding domain" (as generally but not precisely indicated by the stippled regions in the structures above). The FK-506-FKBP complex then binds calcineurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



15

The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506, FK-520, and rapamycin. Modifications in the effector domains of FK-506, FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont *et al.*, 1992, *Journal of Experimental Medicine* 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

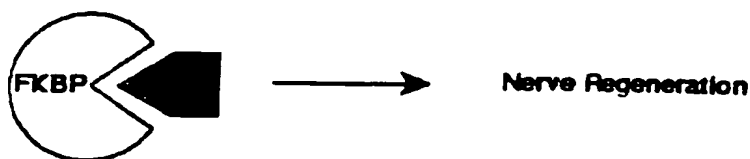
In addition to immunosuppressive activity, FK-520, FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15: 7509-7516; and Steiner *et al.*, 1997, *Proc. National Academy of Science* 94: 2019-2024. Further, the restored central and peripheral neurons appear to be functional.

Compared to protein neurotrophic molecules (BDNF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects. Finally, whereas protein neurotrophins often show effects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

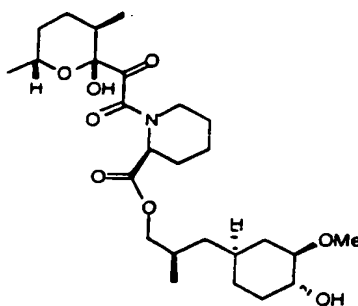
Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine* 3: 421-428.



13



Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne *et al.*, 1993, *Journal of Molecular Biology* 229: 105-124.).  
 Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.

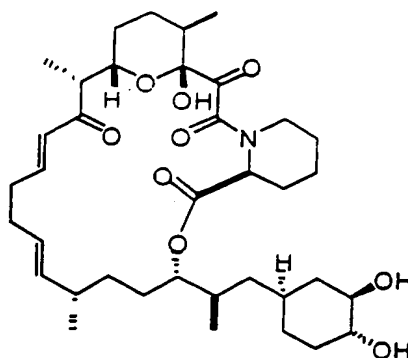


"FKBP binding domain"

There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineurin or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

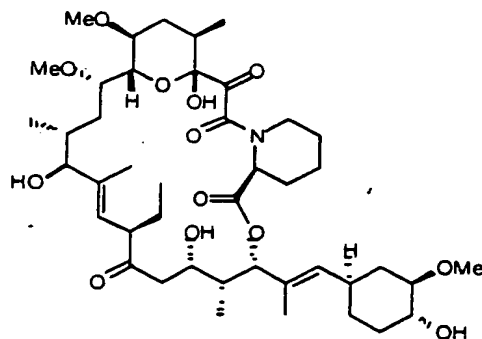
Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr *et al.*, 1996, *The Journal of Antibiotics* 49: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

14

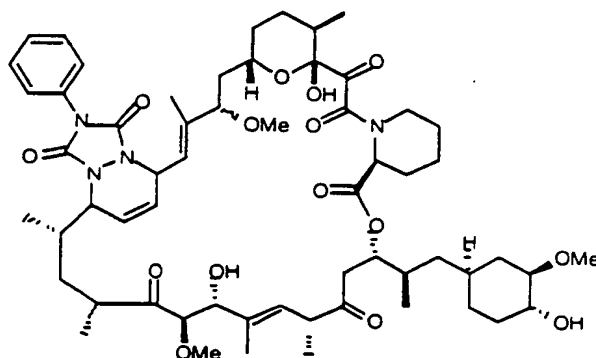


Antascomycin A

Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited, some useful chemically modified analogs exist. The FK-520 analog L-685,818 ( $ED_{50} = 0.7$  nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 ( $IC_{50} = 12.5$  nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications* 192: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).



L-685,818

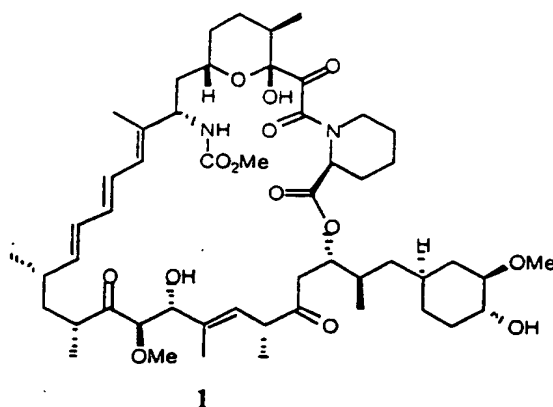


WAY-124,466

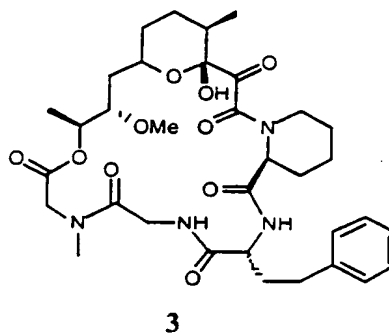
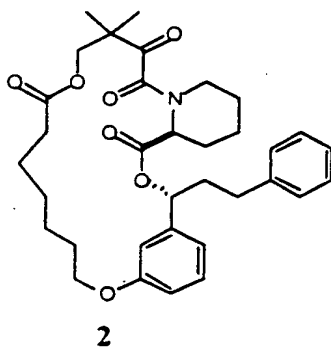
One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology* 2: 471-481). One of the best compounds, 1, below, shows complete

SUBSTITUTE SHEET (RULE 26)

loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.



There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society* 115: 9925-9938); the best analog, 2, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the FKBP binding domain were tethered by a tripeptide to give analog 3, below, which binds to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.



In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand

restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

5 From the above description, two general approaches towards the design of non-immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by  
10 computational methods, and the analogs closely resemble parent molecules that have proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for  
15 production of the numerous compounds needed for such interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

20 The present invention provides useful methods and reagents related to the first approach, but with significant advantages. The invention provides recombinant PKS genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of  
25 which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP.  
30 Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures *via* genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin);

similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

5 Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

10 A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab initio* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated  
15 by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

20 Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical  
25 modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolism and  
30 pharmacokinetics of tacrolimus has been extensively studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%.

(range 5 to 65%). The volume of distribution (VolD) based on plasma is 5 to 65 L per kg of body weight (L/kg), and is much higher than the VolD based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha<sub>1</sub>-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels. Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent *et al.*, 1992, *In vitro* metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, *Arch. Biochem. Biophys.* 294: 454-460; Iwasaki *et al.*, 1993, Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, *Drug Metabolism & Disposition* 21: 971-977; Shiraga *et al.*, 1994, Metabolism of FK-506, a potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, *Biochem. Pharmacol.* 47: 727-735; and Iwasaki *et al.*, 1995, Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506, *Drug Metabolism & Disposition* 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy) compounds undergo cyclizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII,

was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation. Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-506, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-506 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the compounds.

These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and/or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood can vary widely depending on interactions with drugs that induce or inhibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawa US, Rev 4/97, Rec 6/97). Also important are the drugs that induce P450 3A (e.g. Dexamethasone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant adverse effects resulting from the use of FK-506 and are believed to be similar for FK-520. Because these effects appear to occur primarily by the same mechanism as the immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose related and correlates with high blood levels of the drug (Prograf package insert, Fujisawa US, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by the present invention should be more controllable, the incidence of toxicity should be significantly decreased with the 13-desmethoxy analogs. Some reports show that certain FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the *fkfA*, *fkfB*, *fkfC*, and *fkfP* gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the *fkfD* gene product and that is oxidized by the *fkfO* gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the *fkfM* gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded



by the fkbG gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. *ascomyceticus* recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful FK-520-related compound merely as a result of inactivation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCos™ vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of

genomic DNA was partially digested with 4 units of *Sau3A* I for 20 min. in a reaction volume of 1 mL. and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

5       Based on recently published sequence from the FK-506 cluster (Motamedi and Shafiee, 1998, *Eur. J. Biochem.* 256: 528), a probe for the *fkbO* gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids  
10 (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two *EcoRI* fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial  
15 digestion with *Sau3A*I, gel purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced  
20 region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new *fkbM* probe isolated using DNA from ATCC 14891. A probe representing the *fkbP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3  
25 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional  
30 cosmids, pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA, USA. The complete nucleotide sequence of the coding

sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated *fkfB*, *fkfC*, *fkfA*, and *fkfP*. The *fkfB* open reading frame encodes the loading module and the first four extender modules of the PKS. The *fkfC* open reading frame encodes extender modules five and six of the PKS. The *fkfA* open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The *fkfP* open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

	<u>Nucleotides</u>	<u>Gene or Domain</u>
15	complement (412 - 1836)	<i>fkfW</i>
	complement (2020 - 3579)	<i>fkfV</i>
	complement (3969 - 4496)	<i>fkfR2</i>
	complement (4595 - 5488)	<i>fkfR1</i>
	5601 - 6818	<i>fkfE</i>
20	6808 - 8052	<i>fkfF</i>
	8156 - 8824	<i>fkfG</i>
	complement (9122 - 9883)	<i>fkfH</i>
	complement (9894 - 10994)	<i>fkfI</i>
	complement (10987 - 11247)	<i>fkfJ</i>
25	complement (11244 - 12092)	<i>fkfK</i>
	complement (12113 - 13150)	<i>fkfL</i>
	complement (13212 - 23988)	<i>fkfC</i>
	complement (23992 - 46573)	<i>fkfB</i>
	46754 - 47788	<i>fkfO</i>
30	47785 - 52272	<i>fkfP</i>
	52275 - 71465	<i>fkfA</i>
	71462 - 72628	<i>fkfD</i>
	72625 - 73407	<i>fkfM</i>
	complement (73460 - 76202)	<i>fkfN</i>
35	complement (76336 - 77080)	<i>fkfQ</i>
	complement (77076 - 77535)	<i>fkfS</i>
	complement (44974 - 46573)	CoA ligase of loading domain
	complement (43777 - 44629)	ER of loading domain
	complement (43144 - 43660)	ACP of loading domain
40	complement (41842 - 43093)	KS of extender module 1 (KS1)
	complement (40609 - 41842)	AT1
	complement (39442 - 40609)	DH1
	complement (38677 - 39307)	KR1
	complement (38371 - 38581)	ACP1

SUBSTITUTE SHEET (RULE 26)

	complement (37145 - 38296)	KS2
	complement (35749 - 37144)	AT2
	complement (34606 - 35749)	DH2 (inactive)
	complement (33823 - 34480)	KR2
5	complement (33505 - 33715)	ACP2
	complement (32185 - 33439)	KS3
	complement (31018 - 32185)	AT3
	complement (29869 - 31018)	DH3 (inactive)
	complement (29092 - 29740)	KR3
10	complement (28750 - 28960)	ACP3
	complement (27430 - 28684)	KS4
	complement (26146 - 27430)	AT4
	complement (24997 - 26146)	DH4 (inactive)
	complement (24163 - 24373)	ACP4
15	complement (22653 - 23892)	KS5
	complement (21420 - 22653)	AT5
	complement (20241 - 21420)	DH5
	complement (19464 - 20097)	KR5
	complement (19116 - 19326)	ACP5
20	complement (17820 - 19053)	KS6
	complement (16587 - 17820)	AT6
	complement (15438 - 16587)	DH6
	complement (14517 - 15294)	ER6
	complement (13761 - 14394)	KR6
25	complement (13452 - 13662)	ACP6
	52362 - 53576	KS7
	53577 - 54716	AT7
	54717 - 55871	DH7
	56019 - 56819	ER7
30	56943 - 57575	KR7
	57710 - 57920	ACP7
	57990 - 59243	KS8
	59244 - 60398	AT8
	60399 - 61412	DH8 (inactive)
35	61548 - 62180	KR8
	62328 - 62537	ACP8
	62598 - 63854	KS9
	63855 - 65084	AT9
	65085 - 66254	DH9
40	66399 - 67175	ER9
	67299 - 67931	KR9
	68094 - 68303	ACP9
	68397 - 69653	KS10
	69654 - 70985	AT10
45	71064 - 71273	ACP10

1 GATCTCAGGC ATGAAGTCCT CCAGGCGAGG CGCCGAGGTC GTGAACACCT CGCCCTCTCT  
 61 TGTACGGACC ACTTCAGTCA GCGGCGATTG CGGAACCAAG TCATCCGGAA TAAAGGGCGG  
 121 TTACAAGATC CTCACATTGC GCGACCGCCA GCATACGCTG AGTTGCCTCA GAGGCAAACC  
 181 GAAAGGGCGC GGGCGGTCCG CACCAGGGCG GAGTACCGCA CGAGAGTGGC GCACCTGCGC

241 ACCGTACACT CTCTCCCCCG CCGGCGGGAT GCGCGGGGTG ACACGGTTGG GGTCTCTCTG  
 301 ACCCTGAACA CCCGCGCGGT GTGGCGTCGG GGACACTGCC TGGCATCGGC CCGGTGACCG  
 361 TACGGGGGAG GCGTACGGGG GCGGTGGCTC GTGCTACGG CCGCCGGGCG GTCATCCGTG  
 421 GAGACGGGAC TCGGCGAGCA GGGACGCGTG GTGCTACCT GCGGGCGGGA CGACCGTCTG  
 481 GTTCCCGGGG GGGCGGTGGC CCGTGGTGAG CCACTCTCTC AGGGCGGGTA AGGCTGAGCG  
 541 GTGACACGGC AGCAAAGGCC GGAGTCGGTC GGGGAAAGTG TCGACGAGGG CCGTGGTGTG  
 601 CCGTCCCTCC TCGATGCGGT AGTAGCGGTA CCGGCTGGCA GCGCGCTGCC GGACATACCG  
 661 GCGTACACGT CCGAGCCCGG GCGGCAGGCA GCAGCACTGC GAGAGTGCCT GGATGGTCTAT  
 721 CAGCGGCTTC CCGATACGAC CCGTCAACGC GATGCTTTC ACGGCCGCGT GAGCGCCGGA  
 781 GGAGCGGGTG GCGTAGTCGT AGTCGGCATC GCAGCCCGGG ACCGTCCCCG GGGCGCAATA  
 841 CCGTGTGCCG GCTTCTTTCT CCCCATCGAA GCGGGGGTCG AACTCCTCGC GGTAGACGCG  
 901 CTGCGTCAGA TCCAGTAGA CCTCGTGGTG GTACGGCCAC AAGAACTCGG AGTCGGCCGG  
 961 GAGCCCGGCG CCGAGCAGCG CCTCGCGCGC CTGGCCGGCT GCGGGGCGCG CTGCCGCGTA  
 1021 GGTGGGGTAG TCGCGCAGGG CCGCCGGCAG GAAGGTGAAG AGGTTGGGAC CCTCCGCGCG  
 1081 CCACAGGGTG CCTTCCAGT CCACTCCTCC GTGCTACAGC TCGGGATGGT TCTCCAGCTG  
 1141 CCAGCGCAGC AGGTAGCCCG CGTTGGACAT CCCGGTGACC AGGGTGCGCT CGAGCGGGCG  
 1201 GTGGTAGCGC TGGGCGACCG ACGCGCGGGC GGCCCGGGTG AGCTGGGTGA GCGGGGTGTT  
 1261 CCACTCGGCG ACGGCGTCGC CCGCCCGGGA GCCATCACGG TAGAACGCGT GCGCGGTGTT  
 1321 GCGCTTGTCT GTGGCGGCGT AGGCGTAACC GCGGGCGAGC ACCAGTCCGG CGATGGCCCG  
 1381 GTGCTTGGCG TACTGCTCGC GGTACCGGG GGTGCCGGCC ACGACAGGC CACCGTTCCA  
 1441 GCGGTCGGGC AGCCGGATGA CGAACTGGGC GTGCTGGTTC CACCCGTGGT TGGTGTGTTG  
 1501 GGTGGAGGTG TCGGGGAAGT AGCCGTGCTG CTGGATCCCG GGCACCTCCG TGGGAGTGGC  
 1561 CAGGTTCTTG GCGGTACGCG CTGCCAGTC CCGCGGGTCG GTGTGGCCCG TGGCCGCCGT  
 1621 TCCCGCGGTG GTCACTCGT CAGGCGAGTC GCGCTCTGTA CGTGCCCGCG CCGGGACAGC  
 1681 CAGCTGGGAC AGACGGGCGC AGTGACCGTC CCGGGGATCG GGAGCAGGCC GCGCGGCGGC  
 1741 CCGTGAGGGG AGCAGGACCG CCACTGCGGC CAGGCTGAGA GCGCCGAGGC CCGTGCCTCT  
 1801 TCTCGGGGCC CGTCCGACAC CGAGGGGCGA AACCATGAG AGCCTCCAGA CGTGCGGATG  
 1861 GATGACGGAC TGGAGGCTAG GTGCGGCACG GTGGAGACGA ACATGGGTGC GCGCGCCATG  
 1921 ACTGAGGCC CACAGAGGTG GGCCGCGGCC ATGACGGGCG CCGGACCGCG GCGCTCCCG  
 1981 GCGCGGTGCC GCGCGCGCCA CCGGTTCGCG GTCCCGGGT CAGGGACAGG TGTGCTTCG  
 2041 GACGGTGAAG TAGCCGGTCG GCGACTCTTT CAAGGTGGTC GTGACGAAGG TGTGTACAG  
 2101 GCGCATGTTT TGGCCGGAGC CTTGGCGTA GGTGTAACCG GCGCTCGTCG TGGCGCGGCC  
 2161 GCGCTGGACG TGAGCGTAGT TGCCGGCGGT CCAGCAGACG GCGGTGGCAC CCGTCTCTG  
 2221 CCGGGTGACC GCGCCCGAGA GCGGTCCGCG CTTGCCGTCC GCGTCCCGGG CCGCGACCGC  
 2281 GTAGGTGTGC GATGTGCCCC CCGTCAGGCC GGTGTCCGTG TACGACGTCT TGGCGGACGT  
 2341 GGTGATCTGG GCACCGTCGC GGTGGACGGC GTAGTCGGTG GCGCCGTCTG CCGGTTTCCA  
 2401 GGTGAGGCTG ATGGTGGTGT CCGTGGCGCC GGTGGCGGCC AGGCGGACG GAGCGGGCAG  
 2461 CGAACCAGGG TCGGAGGCGG ATCCGCTCAG GCGGAAGAAC TGCGTGATCC AGTACGCGA  
 2521 ACAGATCGAG TCCAGGAAGT AGGCGGCGCC GGTGCTGCCG CACTGCTGTG CTCCGGTGGC  
 2581 GGGATCGACC GGGGTGCCGT GCGCGATGCC CCGCACCGCG TTACCTCCA CCGCCACCGA  
 2641 TCGCTCCGCG GCCAGGTACT CCTCGTGCCG GGTGGAGTTC GGGCGATCA CCGAGGTACG  
 2701 GTCCGGCGTC TGGGACACGC CGTGACACAG GGTCCACTGG TCGCGCAACT CGTCCGCGTT  
 2761 GCGCGGCGCG ACGGTGGTGT CTTGTTCGCC GTGCCAGATG GCCACGCGCG GCCACGGGCC  
 2821 CGACACAGAG GGTAGCCGT CACGACCGC CCGCGCCAC TGGTCCCGCG TCAGGTCCGT  
 2881 CCGGGGGTTC ATGCACAGGT ACGCGCTGCT GACGTCTGTG GCACAGCGGA AGGCGAGGCC  
 2941 GCGGACGACC GCGCCGGCCT GGAAGACGTC CCGATAGGTG GCGAGCATCA CCGACGTCAT  
 3001 GCGACCGCGC GCGGACAGCC CCGTGATGTA GGTGCCCTGG GGGTCCGCGC CGTAGGCGGA  
 3061 GACGGTGTGA GCGGCCATCT GCGCGATCGA CCGGGCTTCG CCCTGGCCCC TCGGGTTCTC  
 3121 GCTGCTCTGG AACCAGTTGA AGCACCTGTT CCGGTTGTTT GACGACGTGG TCTCGGCGAA  
 3181 CACGAGCAGG AAGCCATAGC GGTCCGCGAA TGAGAGCAGG CCGGAGTTGT CCGCGTAGCC  
 3241 CTGGGCGTCC TGGGTGCAAC CGTGACGGGC GAACACCACC CCGGGTCCG CCGGACGGGA  
 3301 CCGGGGCGCG TAGACGTACA GTTTCAGCCG GCGCGGGTTC GTGCCGAAGT CCGGACCTC  
 3361 GGTGAGGTCC GCCTTGGTCA GACCGGGCTT GGCCAGGCC GCGCGGCGGT GGGCGGTCTC  
 3421 CCGCGGGCGC AGCAGGGCCG CTCCGAGTAC GAGGGCCACG ACGGCCACGA GACGGGTGAG  
 3481 CACCCCCCGC CGTCCCGGAC GCGACAACGA CCGGACCGGC GCGGAGGAGG AGAGGGGGA  
 3541 CAGCGGGGTG AGGATTCCCC GGAACGGCGG CCGGTGCATG GCGGCTCCCT CGATGTCTG  
 3601 GGCGGGACAC GGAGGGCTCC CTGACGTCGA TCACTGGGAG CGCCCCGGTG CCGGCGACCG  
 3661 TCGGGGTGT TCAACCGGCA ACGGTATGGC CCGGAGCACC ACACCCCGTA CCGGCGATG  
 3721 TCGCCCCGGA CGGATTGTCT GCGCTTCCG AATCTGATAC CCGGACCGCA CCGACCGCC  
 3781 ACCCGACACG GGTAGGGCGT CATGGTGTCC GACTCGGCCG GTGGGCTTC CCTGCCCTGG

SUBSTITUTE SHEET (RULE 26)

3841 ACGGACCGGG CGTCGGCGGA CCGGGCGTGG GCGGGCTGGG CGGTATGGCG GCGGAGGAGG  
 3901 CCAGCCGCGT GGGGCGGGCG CGCCCAAGTG CAGTACGCGG ACCGTGGGCG GCGGGAGGGG  
 3961 CGGACCGGTC AGTGCAGTCC CGCGGCCCTG CCGGACCGCT CGTCCCAGAC GGGTTCCACC  
 4021 GCGGCGAACC GGGGTCCGTG TCCGCGGGCG TAGACCATCA GTGTCCGCTC GAAGGTGATC  
 5 4081 ACGATGACAC CGTCCTGGTT GTAGCCGATG GTGCGCACGC TGATGATGCC TACGTCAGGT  
 4141 CGGCTGGCGG ACTCCCGGGT GTTCAGGACC TCGGACTGCG AGTAGATGGT STCGCCCTCG  
 4201 AAGACCGGGT TCGGCAGCCT GACCCGGTCC CAGCCGAGGT TGGCCATCAC ATGCTGGGAG  
 4261 ATGTGCGTGA CGCTCTGCCC GGTGACCAGG GCGAGGGTGA AGGTGGAGTC CACCGAGCGG  
 4321 TTGCCCCAGG TGGTGTCCCG CGAGTAGTGG CCGTCAAGT GCGAGCGGCG GGTGTTCTCC  
 10 4381 GTCAGGAGCG TGAGCCAGGA GTTGTGGTTC TCCAGGACCG TCGGGCCCGG GGGGTGGCGG  
 4441 TACACGTGCG CGGTGGTGA GTCTCGAAG TAGCGGCCCT GCGAGCCCTC GACCACAGCG  
 4501 GTGCGGGTGG CGTCTGGTTC CCGGTTCTCA GTCGTCATGG CGCTCATTCT GGAAGTCCC  
 4561 CCGTCCGCTG TGAATGCCG AACCTTCACC GGGCTCATAC GTGCGGCGCA TGAGCCCTGG  
 4621 ACCGTACGTA GTCGTAGAAC CTCGCCACCA CTGGCGCGCG TGGTCTCTCC GCGAGTGTGA  
 15 4681 CCACGCCGAC CGTGCGCCCG GGTGACCGGT CGTCGAGCGG CACGGCGACG GCGTGGTCA  
 4741 CCGGCCCGGA CCGGCTGCCG GTGAGGGGGG CGACGGCCAC ACCGAGGCCG GCGGCGACCA  
 4801 GGGCCCGCAG CGTGCTCAGC TCGGTGCTCT CCAGGACGAC CCGCGGCACG AATCCGGCCG  
 4861 CCGCGCACAG CCGGTCCGTG ATCTGGCGCA GTCCGAAGAC CCGTCCAGT GCCACGAACG  
 4921 CCTCATCGGC CAGCTCCGCG GTCCGCACCC GCGGCGTCT GCGCAGCCGG TGTCCGGGTG  
 20 4981 GGACGAGCAG GCACAGTGCC TCGTCCCGCA GTGGTGTCCA CTCCACATCG TCCCGGCGG  
 5041 GTCGTGGGCT GGTGAGCCC AGGTCCAGCC TGCTGTTGCG GACGTGTCG ACCACGGCGT  
 5101 CCGCGCGCTC GCGCGCAGT TCGAAGGTGG TCGCGGGAGC CAGCGGCGG TACCGGCGA  
 5161 GGAGGTGGGG CACCGCCAG GTGCCGTAGG AGTGCAGGAA ACCCAGTGCC ACGGTGCCG  
 5221 TGTCGGGGTC GATCAGGGCG GTGATGCGCT GCTCGGCGCC GGAGACCTCA CTGATCGCGC  
 25 5281 GCAGGGCGTG GCGCGGAAG ACCTCGCCGT ACTTGTGAG CCGAGCCCG TTCTGGTGCC  
 5341 GGTGGAACAG CCGCACGCC ACTCGTCGCT CCAGCCCGCG GATGGCCCTG GACAGGTGCG  
 5401 GCTGGGAGAT GTTGAGCCGT TCCGCGGTGA TCGTCACGTG CTCGTGCTCG GCCAAGGCCG  
 5461 TGAACCACTG CAACTCCCGT ATCTCCATGC AGGACTATA CGTACCGGGC ATGGTCTTGG  
 5521 CGAGGTTTTC TCATTTTACA GCGGCCGGGC GCGGCCAC AGTGAGTCTT CACCAACAG  
 30 5581 GACCCCATGG GAGGGACCCC ATGTCCGAGC CGCATCCTCG CCTGAACAG GAACGCCCG  
 5641 CCGGGCCCTT GTCCGGTCTG CTCGTGGTTT CTTTGGAGCA GGCCGTGCGC GCTCCGTTCG  
 5701 CCACCCGCCA CCTGGCGGAC CTGGGCGCCC GTGTATCAA GATCGAACGC CCGGCAGCG  
 5761 GCGACCTCGC CCGCGGCTAC GACCGCACGG TCGGTGGCAT GTCCAGCCAC TTCGTCTGCG  
 5821 TGAACCGGGG GAAGGAGAGC GTCCAGTCTG ATGTGCGCTC GCGGAGGGG AACCGGCACC  
 35 5881 TGCACGCCTT GGTGGACCGG CCGCATGTCC TGGTGCAGAA TCTGGACCC GCGCGCGCG  
 5941 GCGGCCTGGC ATCGGCCACC AGGTCTCTCG GCGGAGCCAC CGAGGCTGAT CACCTGCGGA  
 6001 CATATCCGGC TACGGCAGTA CCGGTGCTA CCGCGGACCG CAAGGCGTAC GACCTCTTGG  
 6061 TCCAGTGCGA AGCGGGGCTG GTCTCCATCA CCGGCACCCC CGAGACCCCG TCCAAGGTGG  
 6121 GCCTCTCCAT CGCGGACATC TGTGCGGGGA TGTACGCGTA CTCGGGCATC CTCACGGCCC  
 40 6181 TGCTGAAGCG GGCCCGCACC GGCCGGGGCT CGCAGTTGGA GGTCTCGATG CTCGAAGCCC  
 6241 TCGGTGAATG GATGGGATAC GCCAGTACT ACACGCGCTA CGGCGGCACC GCTCCGGCCC  
 6301 GCGCCGGGCG CAGCCACGCG ACGATCGCCC CTACGCGCCC GTTACCACG GTGACGGGC  
 6361 AGACGATCAA TCTCGGGCTC CAGAACGAGC GGGAGTGGGC TTCTTTCTGC GGTGTCGTGC  
 6421 TACAACGCCC CGGTCTCTGC GACGACCCCG GCTTTTCCGG CAACGCCGAC CCGGTGGCGC  
 45 6481 ACCGCACCGA GCTCGACGCC CTGGTGAGCG AGGTGACGGG CACGCTCACC GCGGAGGAAC  
 6541 TGGTGGCGCG GCTGGAGGAG GCGTCGATCG CCTACGCACG CCAGCGCACC GTGCGGGAGT  
 6601 TCAGCGAACA CCCCCAATG CGTGACCGTG GACGCTGGGC TCGTTCGAC AGCCCGGTGG  
 6661 GTGCGCTGGA GGGCCTGAT CCCCCGGTCA CCTTCCACGG CAGCACCCTG CCGCGGCTGG  
 6721 GCGGGGTCCC GGAGCTGGGC GAGCATACCG AGTCCGTCTT GCGGTGGCTG GCGCGCCCTG  
 50 6781 ACAGCGCCGA CCGCGAAGAG GCGGCGCATG CCGAATGAAC TCACCGAGT CCTGATCCTG  
 6841 GCGCGCGTGT TCCTGCTCGC CGGCGTACGG GGGCTGAACA TGGGCTGCT CCGCTGGTTC  
 6901 GCCACCTTTC TGCTCGGGGT GGTGCGACTC GACCGAACGC CGSACGAGGT GCTGGCGGGT  
 6961 TTCCCGCGGA GCATGTTCTT GGTGCTGGTC GCGGTCACGT TCCTCTTCGG GATCGCCCGC  
 7021 GTCAACGGCA CCGTGGACTG GCTGGTACGT GTCGCGGTGC GGGCGGTGGG GGCCCGGGTG  
 55 7081 GGAGCCGTCC CTTGGGTGCT CTTGGGCTG CCGGCACTGC TGTGCGGAC AGGCGCGGCC  
 7141 TCGCCCGCGG CCGTGGCGAT CTGGGCGCG ATCAGCGTCG CGTTCGCGGT CAGGACCCGC  
 7201 ATCGATCCGC TGTACGCGG ACTGATGGCG GTGAACGGGG CCGCAGCCGG CAGTTTCCGC  
 7261 CCTCCGCGGA TCCTGGGCGG CATCGTCCAC TCGGCGCTGG AGAAGAACCA TGTGCCCTGC  
 7321 AGCGGCGGGC TGCTCTTCGC AGGCACCTTC GCCTTCAACC TGGCGGTGCG CCGGTGTCA  
 60 7381 TGGCTCGTCC TCGGGCGCAG GCGCCTCGAA CCACATGACC TGGACGAGGA CACCGATCCC

	7441	ACGGGAAGGGG	ACCCGGCTTC	CCGCCCCGGG	GCGGAACACG	TGATGACGCT	GACCGCGATG
	7501	GCCGCGCTGG	TGCTGGGAAC	CACGGTCCTC	TCCCTGGACA	CCGGCTTCCT	GGCCCTCACC
	7561	TTGGCGGCGT	TGCTGGCGCT	GCTCTTCCCG	CGCACCTCCC	AGCAGGCCAC	CAAGGAGATG
5	7621	GCCTGGCCCG	TGGTGTGCTG	GGTATGCGGG	ATCGTGACCT	ACGTGCGCCT	GCTCCAGGAG
	7661	CTGGGCAATG	TGGACTCCCT	GGGGAAGATG	ATCGCGGCGA	TCGGCACCCC	GCTGCTGGCC
	7741	GCCCTGTGTA	TCTGCTACGT	GGGCGGTGTC	GTCTCGGCCT	TCGCCTCGAC	CACCGGGATC
	7801	CTCGGTGCCC	TGATGCCGCT	GTCCGAGCCG	TTCCTGAAGT	CCGGTGCCAT	CGGGACGACC
	7861	GGCATGCTGA	TGGCCCTGGC	GGCCGCGGCG	ACCGTGGTGG	ACGCGAGTCC	CTTCTCCACC
10	7921	AATGGTGTCT	TGGTGGTGCG	CAACGCTCCC	GAGCGGCTGC	GGCCCGGCGT	GTACCAGGGG
	7981	TTGCTGTGCT	GGGCGCCCGG	GGTGTGCGCA	CTGGCTCCCG	CGGCCGCTCG	GGCGGCTTTC
	8041	GTGGTGGCGT	GAGCGCAGCG	GAGCGGGAAT	CCCTTGAGAG	CCGTTTCCCG	TGCTGTGTCT
	8101	CTGACGTAGC	GTCAAGTCCA	CGTGCCGGGC	GGGCAGTACG	CCTAGCATGT	CGGGCATGGC
	8161	TAATCAGATA	ACCCTGTCCG	ACACGCTGCT	CGCTTACGTA	CGGAAGGTGT	CCCTGCGCGA
15	8221	TSACGAGGTG	CTGAGCCGGC	TGCGCGCGCA	GACGGCCGAG	CTGCCGGGCG	GTGGCGTACT
	8281	GCCGGTGCAG	GCCGAGGAGG	GACAGTTCCT	CGAGTTCCTG	GTGCGGTTGA	CCGCGCGCGG
	8341	TCAGGTGCTG	GAGATCGGGA	CGTACACCGG	CTACAGCACG	CTCTGCCTGG	CCCGCGGATT
	8401	GGCGCCCGGG	GGCCGTGTGG	TGACGTGCGA	TGTCATGCCG	AAGTGGCCCC	AGGTGGGCGA
	8461	GCGGTACTGG	GAGGAGGCCG	GGGTTGCCGA	CCGGATCGAC	GTCCGGATCG	CGCACGCCCG
20	8521	GACCGTCCTC	ACCGGGCTGC	TCGACGAGGC	GGGCGCGGGG	CCGGAGTCGT	TCGACATGGT
	8581	GTTCATCGAC	GCCGACAAGG	CCGGCTACCC	CGCCTACTAC	GAGGCGGCGC	TGCCGCTGGT
	8641	ACGCCGCGGC	GGGCTGATCG	TCGTGACAAA	CACGCTGTTC	TTCGGCCGGG	TGGCCGACGA
	8701	AGCGGTGCAG	GACCCGGACA	CGGTGCGGGT	ACGCGAACTC	AACGCGGCAC	TGCGCGACGA
	8761	CGACCGGGTG	GACCTGGCGA	TGCTGACGAC	GGCCGACGGC	GTCACCTCTC	TGCGGAAACC
25	8821	GTGACCGGGG	CGATGTCGGC	GCGCGTCAAG	TCGACGCTCG	TCGGCGCGGG	CCTCGCGGAG
	8881	GGCTCCAGAT	GCAGGCGTTC	GACGCCGGCG	GCGGAAGCGC	CCGCCACCTC	GGACACGCAG
	8941	GGGCGATCGG	AGTCCGCGAA	GCCCGCGAAC	CGGTAGGCGA	TCTCCATCAT	GCGGTTGCGG
	9001	TCCGTACGCC	GGAAGTCCGC	CACCAGGTGC	GCCCCCGCSC	GGGCGCCCTG	GTCCGTGAGC
	9061	CAGTTCAGGA	TCGTGCGACC	GGCACCGAAC	GACACGACCC	GGCAGGACGT	GGCGAGCAGT
30	9121	TTCAGGTGCC	ACGTGCGACG	CTTCTTCTCC	AGCAGGATGA	TGCCGACGGC	GCCGTGCGGG
	9181	CCGAAGCGGT	CGCCCATGGT	GACGACGAGG	ACCTCATGGG	CGGGATCGGT	GAGCACGCGC
	9241	GCAAGTCCGC	GTCCGAGTAG	TGCACGCCGG	TCGCGTTTAT	CTGGCTGGTC	CGCAGCGTCA
	9301	GTTCTCTGAC	GCGGCTGAGT	TCCTCTCTCC	CCGCGGGTGC	GATCGTTCATG	GACAGGTCGA
	9361	GCGAGCGCAG	GAAGTCCTCG	TCGGGACCGG	AGTACGCCTC	CCGGGCTTGG	TCGCGCGCGA
35	9421	AACCCGCTTG	GTACATCAGG	CGGCGCCGAC	GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
	9481	ACTCCGGCAG	CGACAGGAGC	GTGGCCGCTT	GCTCGGCCGG	GTAGCACCAG	ACCTCGGGCA
	9541	GGTGGAAACG	CACCTCGGCA	CGCTCGGCGG	GCTGGTCTGT	GATGAACGCG	ATCGTGGTCT
	9601	GTGCGAAGTT	CAGTCCGTG	GCGATCTCGG	GGACGGACTG	CGACTTCGGC	CCCCATCCGA
	9661	TGCGGGCCAG	CACGAAGTAC	TCCGCCACAC	CGAGGCGTTC	CAGACGCTCC	CACGCGAGGT
40	9721	CGTGGTCTGT	CTTGCTCGCC	ACCGCCTGGA	GGATGCCGCG	GTGCTCGAGC	GTGGTGATCA
	9781	CCTCGCGGAT	CTCGTCCGTG	AGGACCACCT	CGTCTCTCTC	CAGCACGGTG	CCCCGCCACA
	9841	AGGTGTTGTC	CAGGTCCGAG	ACCAGACACT	TGACAATGGT	CATGGCTGTC	CTCTCAAGCC
	9901	GGGAGCGCCA	GCGCGTGCTG	GGCCAGCATC	ACCCGGCACA	TCTCGCTGCT	GCCCTCGATG
	9961	ATCTCCATGA	GCTTGGCGTC	GCGGTACGCC	CGTTCCAGCA	CGTGTCCCTC	TCTCGCGCCT
45	10021	GCCGACGCGA	GCACCTGTGC	GCGGTGCGCG	GCCCGGCGCG	CGGCTCGTTC	GGCGGCGAGC
	10081	TGCTTGGCCA	GGATCGTTCG	GGGCACCATC	TCGGGCGAGC	CCTCGTCCCA	GTGGTCTGCT
	10141	GCGTACTCGC	ACACGCGGGC	CGCGATCTGC	TCCGCGGTCC	ACAGGTGCGC	GATGTGCCCG
	10201	GCGACGAGTT	GGTGGTCCGC	GAGCGGCGCG	CCGAACTGCT	CCCGGGTCCG	GGCGTGGGCC
	10261	ACCGCGGCGG	TGCGGCGAGG	CCGCGAGGATC	CCGACGCGAG	CCGAGGCGAC	CGACTTGCGC
50	10321	CCGTAGGCGA	GTGACGCCGC	GACCAGCATC	GGCAGTGACG	CGCCGGAGCC	GGCCAGGACC
	10381	GCGCCGCGCG	GCACACGCAC	CTGGTCCAGG	TGCAGATCGG	CGTGGCCGGC	GGCGCGGCAG
	10441	CCGGACGCGT	TGCGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTCGGCGGG	CACGACCACC
	10501	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACCA	GGTGGTCCGC	GTAGGCGGGC
	10561	GCACTCGTCC	AGACCTTGTC	GCCGTGCGAG	ACAGCGGTGT	CCCCGTGCGG	CCGAACCCGC
55	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGCCGAC	GGCCGCGAGT
	10681	TTCCCGCTGG	TCAGTCTCTT	CAGGAAGGTC	GCCCCGCTGAC	CGGCGTCCGC	GAGCCGCTGC
	10741	ACGGTCCACG	CGGCCATGCC	CTGCGACGTC	ATGACACTGC	GCAGCGAATC	GCAGAGGCTG
	10801	CCGACGTGTC	CGGTGAACTC	GCCGTTCTCC	CGGCTGCCGA	GTCCGAGACC	CCGCTGCTCC
	10861	CCGCGCACTT	CCGCGCAGAG	CAGGTCCTCG	GCGCGGAGCC	GGACCGAGCC	GTCGCGCGGC
60	10921	AGTTCGCGCG	ACGTGTCCCA	CTCGGCGGCC	CGGTACCGCA	CAAGGTGCGT	CAGCAGCGCG
	10981	TCACGCTCAG	GCATCGACGG	CCCGCAGCCG	GTGGACGAGT	GCGACCATGG	ACTCGACGGT

SUBSTITUTE SHEET (RULE 26)

11041 ACGGAAGTTC GCGAGCTGGA GGTCCGGGGC GGGGATCGTG ACGTCGAACG TCTTCTCCAG  
 11101 GTACACGACC AGTTCCATCG CGAACAGCGA CGTGAGGCGG CCTCCGCGA ACAGGTCCGG  
 11161 GTCCACGGGC CAGTCCGACC TGGTCTTCGT CTTGAGGAAC GCGACCAACG CGTGCGCGAC  
 11221 GGGGTCTGTC TTGACGGGTG CGGTCTATGA AACACCTTCT CGTATTCGTA GAAGCCCCCG  
 5 11281 CCGGTCTTCC GGCCGTGGTG TCCCTCGCGG ACCTTGCCCA GCAGCAGGTC ACAGGGGGCG  
 11341 CTGGGCTCGT CGCCGGTGCG TTTGTGCAGC ACCACAGCG CGTCGACGAG GTTCTCGATG  
 11401 CCGATCAGGT CCGCGGTGCG CAGCGGGCCG GTCGGATGGC CGAGGCACCC CGTCATGAGC  
 11461 GCGTCGACGT CCTCGACGGA CGCGGTGGCC TCCTGCACGA TCCGCGCCCG GTCGTTGATC  
 11521 ATCGGGTGGA GCAGCCGGCT CGTGCAGGAG CGGGGCGCGT CCGGACGAC CATCGGCTTG  
 10 11581 CCGCCGACGG CCGCGAGCAG GTCCCGGGCG GCGGCCATGG CCTTCTCACC GGTCCGGGT  
 11641 CCGCGGATCA CCTCGACCGT CGGGATCAGG TACGACGGGT TCATGAAGTG CGTGCCGAGC  
 11701 AGGTCTTCGG GCCGGGGCCAC GGAGTGGGCG AGTTCGTCAA CCGGGATCGA CGACGTGTTT  
 11761 GTGATGACCG GGATACCGGG CGCCGCTGCG GAGACCGTGG CGAGTACCTC CGCCTTGACC  
 11821 TCGGCGTCCT CGACGACGGC CTCGATCACC GCGGTGGCCG TACCGATCGC GGGCAGCGCG  
 15 11881 GACGTGGCCG TCCGAGCAG ACCGGGCTCG GCCTCGCGG GCCCGGCCAC GAGTTGTGCC  
 11941 GTCCGCAAGT CCGGTGGCGAT CCGCGCCCGC CCGCGCGTAA GGATCTCCTC GGACGTGTCG  
 12001 ACGAGTGTCA CCGGGACGCC GTGGCGCAGC GCGAGCGTGG TGATGCCGGT GCCCATCACT  
 12061 CCGCGCCCGA GCACGATCAG CTGGTGGTCC ACGCTGTTTC CTCCTCCCGG GTCCACCATG  
 12121 GCAGCGAGTA CGGGTGCGAG ACGTCTTCGG GGGTCGACCC GATCGCGTCC TTGCGGCCGA  
 20 12181 GGCCGAGTTC GTCCGGCAAG CCGAGCAGCA CGTCGAACGC GATGTGGTCG GCGAACGCGC  
 12241 TGCCCGTCGA GTCGAGGACG CTCAGGCTGT CCGGTGGTCC GCGCGGGTG TCCGGTGCCG  
 12301 GCGACAGGGC CGCCAGCAG CGCGGTCCGG CAGTTGCTGG TACTCGCCCT  
 12361 CGGCGCGGGC CTGCCCCGGA TGGTCGACGC AGATGAACGC GTCGTCGAGC AGGTCTTTCG  
 25 12421 GCAATTCGGT CTTGCCCGGC TCGTCGGCGC CGATGGCGTT CACATGCAGG TCGCGCAGCC  
 12481 GCGGCTCGGC GGGCAGCACC GGCCCTTTTC CCGAGGGCAC CGAGGTGACG GTGGACAGGA  
 12541 CATCCGCGGC GCGCGCGGCC TCCGCCCGAT CGGTCACTT GACCGGCAGT CCGAGGAACG  
 12601 CGATGCGGTC CGCGAACGAC GCGCGCTGSC CCGGGTCCGT GTCGCTGACC AGGATCCGCT  
 12661 GATGGGCGAG GACCTGTCTG AGCGCTGCG CCTGGGTCAC CGCTGTGCG CCCGCGCCGA  
 30 12721 TCAGCGTGAG CGTGGCGCTG TCGGACCGGG CCGAGAGCCG GCTCGCAGC GCGCGACCG  
 12781 CGCCGGTCCG CATCGCGGTG ATCAGCCTG CGTCGGCGAG GCGCGTCAGA CTGCCGTGT  
 12841 CGTCGTCGAG GCGCGACATC GTGCCGACGA TCGTCGGCAG CCGGAAGCGC GGATAGTTGT  
 12901 GCGGACTGTA CGAAACCGTC TTCATGGTCA CGCCGACACC GGGGACCCGG TACGGCATGA  
 12961 ACTCGATGAC GCCGGGAATG TCGCCGCCGC GGACGAATCC GGTACGCGGC GCGCCTCGG  
 35 13021 CGAACTCGCC GCGGCCGAGC GCGGCGAACC CGTCGTGACG CTCGCTGATC AGCCGGTCCA  
 13081 TCATCAGTC GCGGCCGATC ACGGAGAGA TCCGCTTGAT GTCACGTTGC CGCAGGACCC  
 13141 TGGTCTGCAT GTGTACCTC CTTTCTGTG CGGAGCTGT CTTGGTGGTG CCGCTCGGG  
 13201 CGGCTTCCGT TCTCATCGCA GCTCCCTGTC GATGAGGTG AAAATCTCGT CCGCGGTGCG  
 13261 GTCCGCGGAC AGCACGCCGG CCGGCGTGT CCGGCGGGTC TCCCGCCGCC AGCGTTGAG  
 13321 CAGGGCGTCC AGCCGGGTTT CGATCGCGTC CGCCTGGCG GCGCCCGGGT CGACACCGGC  
 40 13381 AACGAGTGCT TCCAGCCGGT CGAGCTGCGC GAGCACCACG GTCACCGGGT CGTCCGGGGA  
 13441 CAGCAGTTCA CCGATGCGGT CGGCGAGTGC GCGCGGCGAC GGGTAGTCGA AGACGAGCGT  
 13501 GCGCGACAGT CGCAGACCGG TCGCTCGTT GAGGCCGTTG CGCAGCTGCA CCGCATGAG  
 13561 CGAGTCCACA CCGAGTTCCC GGAACGCCGC GTCCTCCGG ATGTCCTCCG GGTGCGGCTG  
 45 13621 GCGCAGGACG GCGCTGCCT TCTGCCGAC GAGGGCGAGC AGGTGCGTGG GCGTTCTTG  
 13681 CTCGTTGCG GCGCTCCGGC GGGCCGACCG CTTGGGCGCG CCACGCGACA GCGGGAGGTC  
 13741 CCGCGGCGAG TCGCCCGCCA CGGCGACGAC ACTGCCCGTT CCGGTGTGGA CCGCGGCGTC  
 13801 GTACATGCGC ATGCCCTGTT CCGCGGTGAG CGCGCTCGCC CCACCTTGC GCATACGGCG  
 13861 CCGGTCCGGC TCGTCAAGT CCGCGGTGAG GCCACTCGCC TGGTCCACA GCGCCACGCG  
 50 13921 GATCGACAGC CTTGGCAGCC CTTGTGCA GCGGTGTTG GCGAGCGCGT CGAGGAACGC  
 13981 GTTCCCGGCC GCGTAGTTGC CTTGACCGGG GGTGCCAGC ACACCGGCCG CCGACGAGTA  
 14041 GACGACGAAT GCGGCGAGGT CCGGTGTCGCG GGTGAGCCGG TGCAGGTGCC AGGCGGCGTC  
 14101 GGCCTTGCGT TTGAGGACGG TGTGATGCG GTCCGGGGTG AGGTTGTGCA GCAGGGCGTC  
 14161 GTCGAGGGTT CCGGCGGTGT GGAAGACGCG GGTGAGGGGT TGAGGGATGT GGGCGAGGGT  
 14221 GGTGGCGAGT TGGTGGGGGT CCGCGACGTC GCAGGGGAGG TGGGTGCCGG GGGTGGTGTC  
 55 14281 GGGGGGTGG GTGCGGGAGA GGAGGTAGGT GTGGGGGTGG TTCAGGTGGC GGGCGAGGAT  
 14341 CCGCGGAGG GTGCGGAGC CGCCGCTGAT GACGACGGCC CCTCGGGGT CCAGCGGCCG  
 14401 CCGGACCGTG AGGACGATCT TGCCGCTGTG TCGCCCGCG CCGATGGTCG CCAGCGCTC  
 14461 GCGGACCTGC CGCATGTCGT GCACCGTCA CCGCAGCGGG TGCAGCACAC CGCGCGCGAA  
 14521 CAGGCCGAGC AGCTCCGCGA TGATCTCCTT GAGCCGGTGC GGCCCGCGT CCATCAGGTC  
 60 14581 GAACGGTCCG TGGACGGCGT GCCGGATGTC CTTCTTCCCC ATCTCGATGA ACCGGCCACC



	14641	CGGCGCGAGC	AGGCCGACGG	ACGCGTCGAG	GAGTTCACCG	GTGAGCGAGT	TGAGCACGAC
	14701	GTCGACCGGC	GGGAACGCGT	CGGCGAACGC	GGTGCTGCGG	GAATCGGCCA	GATGCGCTCC
	14761	GTCCAGGTCC	ACCAGATGGC	GCTTCGCGGC	GCTGGTGGTC	GCGTACACCT	CGCGCGCCAC
5	14821	GTCCCGCGCG	ATCTGCCGGG	CGGCGGAACC	GACACCGCCG	GTGGCGCGCT	GGATCAAGAC
	14881	CTTCTCGCCG	GGGCGCAGCC	CGGCGAGGTC	GACCAGGCCG	TACCACGCGG	TCGCGAACGC
	14941	GGTCATCAGC	GACGCCGCCT	GCGGGAACGT	CCAGCCGTCC	GGCATCCGGC	CGAGCATCCG
	15001	GTGGTCCGGC	ATGACCGTGG	GGCCGAAGCC	GGTGCCGACG	AGGCCGAAGA	CGCGGTCCGC
	15061	GCGTGCCAGA	CCGGAGACGT	CGGCGCCGGT	CTCCAGGACG	ATGCCCGCGG	CCTCGCCGCC
10	15121	GAGCACGCCG	TGACCGGGGT	AGGTGCGCAG	CGCGATCAGC	ACATCGCGCA	AGTTGAGGCC
	15181	CGCGCGACGC	ACACCGATCC	GGACCTCGGC	CGGGGCGAGG	GGGCGCCGGG	GCTCCGCCGA
	15241	GTGCGCCCGC	GTGAGGCCGT	CGAGGGTGCC	CGTCCGCGCC	GGCCGGATCA	GCCACGTGTC
	15301	GCTGTCCCGC	ACGGTGAGCG	GCTCCGGCAC	CCGGGTGAGG	CGGGCCGCCT	CGAACCAGCC
	15361	GCGCGCGACG	CGCAGACGCG	GCTCGCCGAG	TGCGACGCGC	ATGCGCTGCT	GCTCGGGGGC
15	15421	GAGCGTGACG	CCGGACTCGG	TCTCGACGTG	GACGAACCGG	CCGGGCTGCT	CGGCCCTGGC
	15481	GCGCGCGACG	AGTCCGGCCG	CGGCGCCGGT	GGCGAGGCCG	GCGGTGGTGT	GCACGAGCAG
	15541	ATCCCCGCGC	GAGCCGGTCA	GGGCGCGTAG	CAGCCGGGTG	GTGAGCGCAC	GCGTCTCGGC
	15601	CACCGGGTCG	TCGCCATCAG	CGGCAGGCAA	CGTGATGACG	TCCACGTGCG	TCGCGGGGAC
	15661	ATCCGTGGGT	GCGGCGACCT	CGATCCAGGT	GAGACGCATC	AGGCCGGTGC	CGACGGGTGG
20	15721	GGACAGCGGG	CGGGTGCGGA	CCGTCCGGAT	CTCGGCGACG	AGTTGGCCGG	CGGAGTCGGC
	15781	GACGCGCAGA	CTCAGCTCGT	CGCCGTACAG	AGTGATCAGC	GCTCGGAGCA	TGGCCGAGCC
	15841	CGTGGCGACG	AACCGGGGCC	CCTTCCAGGC	GAACGGCAGA	CCCGCAGCGC	TGTCGTCCGG
	15901	CGTGGTGAGG	GCGACGGCGT	GCGCGCCGCG	GTGAGCAGC	GCGGATGCA	CACCGAAACC
	15961	GTCCGCGCTG	GCGGCCTGCT	CGTCCGGGAG	GCGCACCTCG	GCATACACGG	TGTACCATC
25	16021	ACGCCAGGCA	GCCCCGAAAC	CCTGGAACGC	CGACCCGTAC	TCATAACCGG	CATCCCGCAG
	16081	TTCGTATAG	AACCCCGAGA	CGTCGACGGC	CACGGCCGTG	ACCGGCGGCC	ACTGCGAGAA
	16141	CGGCTCCACA	CCGACAACAC	CGGGGGTGTC	GGGGGTGTG	GGGGTCAGGG	TGCCGTGGC
	16201	GTGCCGGGTC	CAGCTGCCCC	TGCCCTCGGT	ACGCGCGTGG	ACGGTCAACG	GCCGCCGTCC
	16261	GGCTTCATCA	GCCCCCTTCCA	CGGTCAACCA	CACATCCACC	GCTGCGGTCA	CGGGCACACC
30	16321	AAGGGGGGAT	TCGATGACCA	GCTCGTCCAC	TATCCCGCAA	CCGTCTCTGT	CACCGGCCCG
	16381	GATGACCAGC	TCCACAAACG	CCGTACCCGG	CAGCAGGACC	GTGCCCCGCA	CCCGCTGATC
	16441	AGCCAGCCAG	GGGTGAGTGC	GCAATGAGAT	CCGGCCAGTG	AGAACAACAC	CACCATCGTC
	16501	GGCGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CGGATGCGCC	GCACCCGTCA	ACCCCGCCGC
	16561	CGACAGATCG	GTGGCACCCG	CCGCCTCCAG	CCAGTACCGC	CTGTGCTCGA	ACGCGTACGT
35	16621	GGGCAGATCC	AGCAGCCGTC	CCGGCACCCG	TTCGACCACC	GTGTCCCAGT	CCACTGCCGT
	16681	GCCCAGGGTC	CACGCCGTGC	CCAACGCCGT	CAGCCACCGC	TCCCAGCCCG	CGTCACCGGT
	16741	CCGCAACGAC	GCCACCGTGT	GCCACGTCTC	CATCGCCGGC	AGCAGCACCG	GATGGCGACT
	16801	GCACTCCACG	AACACCGACC	CATCCAGCTC	CGCCACCGCC	GCGTCCAACG	CCACCGGACG
	16861	ACGCAGATTC	CGGTACCACT	ACCCCTCATC	CACCGGCTCC	GTCACCCAGG	CGCTGTCCAC
40	16921	GGTCGACCAC	CACGCCACCG	ACGCGGCCTT	CCCTGCCACC	CCCTCCAGTA	CCTTGGCCAG
	16981	TTCATCCTCG	ATGGCTTCCA	CGTGGGGCGT	GTGGGAGGCG	TAGTCGACCG	CGATACGACG
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
	17101	CGCCACCAAC	GTCGAAGCCG	GGCGGTTACG	CGCCGCGATC	CACACACCTT	CGACCAAGCC
	17161	GACCTCACCG	GCCGGCAACG	CCACCGAAGC	CATCGCTCCC	CGCCCGGCCA	GTCGCGCCGC
45	17221	GATGACCTGA	CTGCGCAATG	CCACCACGCG	GGCGGCGTCC	TCGAGGCTGA	GGGCTCCGGC
	17281	CACGCACGCC	GCCGCGATCT	CGCCCTGGGA	GTGTCCGATC	ACCGCGTCCG	GCACGACCCC
	17341	ATGCGCCTGC	CACAGCGCGG	CCAGGCTCAC	CSCGACCGCC	CAGCTGGCCG	GCTGGACCAC
	17401	CTCCACCCGC	TCCGCCACAT	CCGGCCGCGC	CAACATCTCC	CGCACATCCC	AGCCCGTGTG
	17461	CGGCAGCAAC	GCCTGAGCGC	ACTCCTCCAT	ACGCGCGGCG	AACACCGCGG	AGTGGGCCAT
50	17521	GAGTTCCACG	CCCATGCCGA	CCCACTGGGC	GCCCTGGCCG	GGGAAGACGA	ACACCGTAGC
	17581	CGGCTGGTCC	ACCGCCACAC	CCGTACCCCG	GGCATCAGCC	AGCAGCACCG	CACGGTGACC
	17641	GAAGACAGCA	CGCTCCCGCA	CCAACCCCTG	CGCGACCGCG	GCCACATCCA	CACCAACCCC
	17701	GCGCAGATAC	CCCTCCAGCC	GCTCCACCTG	CCCCCGCAGA	CTCACCTCAC	CACGAGCCGA
	17761	CACCGGCAAC	GGCACCAACC	CGTCAACAAC	CGACTCCCCA	CGCGACGGCC	CAGGAACACC
55	17821	CTCAAGGATC	ACGTGCGCGT	TCGTACCGCT	CACCCCGAAC	GACGACACAC	CCGCATGCGG
	17881	TGCCCCGATC	GACTCGGGCC	ACGGCCTCGC	CTCGGTGAGC	AGCTCCACCG	CACCGGCCGA
	17941	CCAGTCCACA	TGCGACGACG	GCTCGTCCAC	ATGCAGCGTC	TTCGGCGCGA	TCCGCTACCG
	18001	CATCGCCATG	ACCATCTTGA	TCACCCCGGC	GACACCCGCC	GCCGCTGCGG	CATGACCGAT
	18061	GTTGCACTTC	AACGAACCCA	GCAGCAGCGG	AACCTCACGC	TCCTGCCCGT	ACGTGCGCAG
60	18121	AATGGCCTGC	GCCTCGATGG	GATCGCCCGC	CGTCGTCCCC	GTCCCGTGCG	CCTCCACCAC
	18181	GTCCACATCG	GCGGCGCGCA	GTCCGGCGTT	CACCAACGCC	TGCTGSATGA	CACGCTGCTG

SUBSTITUTE SHEET (RULE 26)

18241 GGACGGGCGG TTGGGGGCGG ACAGCCCGTT GGAGGCACCG TCCTGGTTCA CCGCCGACCC  
 18301 GCGGACGACC GCGAGAACGG TGTGTCCGTT GCGCTCGGCG TCGGAGAGCC GCTCCAGCAC  
 18361 AAGAACGCGG GCGCCCTCCG CCCAGCCGGT GCGGTTGGCG GCGTCCGCGA ACGCGCGGCA  
 18421 GCGGCCGCTG GGGGAGAGTC CGCCCTGCTG CTGGAATTCC ACGAACCCGG TCGGGGTCTG  
 5 18481 CATGACGGTG ACACCGCCGA CCAGCGCCAG CGAGCACTCC CCGTGGCGCA GTGCGTGCCC  
 18541 GGCCTGGTGC AGCGCGACCA GCGACGACGA GCACGCCGTG TCCACCGTGA ACGCCGGTCC  
 18601 CTGGAGCCCA TAGAAGTACG AGATCCGGCC GGTGAGCACG CTGGGCTGCA TGCCGATCGA  
 18661 GCCGAACCCG TCCAGGTCCG CGCCGACGCG GTACCCGTAC GAGAAGGCGC CCATGAACAC  
 18721 GCGGGTCTCG CTGCCGCGCA GTGTGCCCGG CACGATGCCC GCGCTCTCGA CGCCTCCCA  
 10 18781 TGTCGTTTCC AGCAGGATCC GCTGCTGGGG GTCCATGGCC CGTGCTCAC GGGGGCTGAT  
 18841 GCCGAAGAAC GCGGCATCGA AGCCGGCGGC GTCGGAGAGG AAGCCGCCGC GGTCCGTGTC  
 18901 CGATCCGCGG GTGAGGCCGG ACGGGTCCCA GCCACGGTCG GCCGGGAAGC CCGTGACCGC  
 18961 GTCGCCGCCA CTGTCCACCA TGCGCCACAG GTCGTCCGGC GAGGTGACGC CGCCCGGCAG  
 19021 TCGCAGGCGG ATGCCACGA TGGCCAGCGG TTCGTACCGG GTCGCGGCGG CTGTGGGAAC  
 15 19081 AGCGACCGGT GCGGCACCA CAGCCAGAGC CTCGTCCAAC CGCGACCGA TGGCCGCGG  
 19141 CGTCGGGTAG TCGAAGACAA GCGTGGCGGG CAGTCGGACA CCGGTGCGCG CGCGAGTCTG  
 19201 GTTCCGCACT TCGACGGCGG TCAGCGAGTC GATACCCAGT TCCTTGAAGG CCGCGTCCGC  
 19261 GGACACGTCC GCGGCGTCCG CGTGCCGAG CACCGCCGCC GCGTTGTGCG GGACAGTGC  
 19321 CAGCAGCGCG GTGTCCCGCT CAGCGCCGGA CATGGTGCCG AGCCGGTCCG CGAGCGGAAC  
 20 19381 GCGGGTGGCC GCCGCCGGGC GCGATACGGC GCGGCGCAGA TCGGCGAAAA GCGGCGATGT  
 19441 GGCGCGGCTG AGGTCCATCG TGCCCGCCAC GCGGAACGCG GTGCGGTTTC CGGCCGCGGC  
 19501 TTCCAGCAGG CGCATGCCCC CACCGCCGGA CATGGGGCGG AAACCGGCTC AGGGACACG  
 19561 GGTGCGGTTG GTGCCGCTCA TGCTGCCGGT GAGTCCGCTG TCATCGGCCC AGAGCCCCCA  
 19621 GGCCAGCGAC AGCGCGGGCA GTCCTTCGGC ATGGCGCAGC GTCGCGAGTC CGTCGAGGAA  
 25 19681 CCCGTTCCGC GCCGAGTAGT TGCCCTGGCC GCGGCCGCC ATGATGCCCG CGACGGACGA  
 19741 GTAGAGGACG AACGAGCGCA GGTCCGCGTC CCGGGTCAGC TCGTGCAAGT GCCAGGCGCC  
 19801 GTCGGCTTTG GGGCGCAGTG TGGTGGCGAG CCGCTCCGGG GTGAGTGCCG TGGTCACGCC  
 19861 GTCCGTCAGC ACGGCTGCCG TGTGGAAGAC CGCCGTGAGC GGCCTGCCG CGCGCGCAG  
 19921 CGCGGCGGCG AGCTGGTCCC GGTGCGGAC GTACAGCGG ATGTGGACAC CGGAGTGTG  
 30 19981 CGCCGGCGGT TCGCTGCGCG ACAGCAACAG GAGGTGGCGG GCGCCATGCT CGGCGACGAG  
 20041 ATGCCGGGCG AGGAGACCTG CCAGCACACC CGAGCCGCGG GTGATGACCA CCGTGCCGTC  
 20101 CCGGTTCGAGC AGCGGTTCCG GCGTTTCCGC GCGCGCCGTG CCGGTGAACC GCGGCGCTTC  
 20161 GTACCGGCCG TCGGTGACGC GGACGTACGG CTCGGCCAGT GTCGTGCGCG CGGCCAGCGC  
 20221 CTCGATGGGG GTGTGCGTGC CGGTCTCCAC CAGCACGAAC CGGCCCGGGT GCTCGGCTG  
 35 20281 CTCGGACCGG ACGAGGCGCG CGACCGCTCC TCCGACCGGT CCCGCTCGA TCCGGACGAC  
 20341 GAGGGTGGTC TCCGCAAGGC CGTCTCGGC GATCACCCGG TGCAGTCTCG CGAGCACGAA  
 20401 CTCGGTGAGC CGGTACGTCT CGTCGAGGAC ATCCGCGCCC GGTTCGGGA GCGCGGAGAC  
 20461 GATGTGGACC GCGTCCGCG GACCGGGCCC GGGAGTGGGC AGCTCGGTCC AGGAGAGGCC  
 20521 GTACAAGGAG TTCCGTACGA CCGCGGCGTC GCGGTCGACG TTCACCGGTC GCGCGTCTAG  
 40 20581 CGCGGCGACG GTCACCACCG GTTGGCCGAC CCGGTCCGTC GCATGCACGG CAGCGCCGTC  
 20641 CGGGCCCTGA GTGATCGTGA CGCGCAGCGT GGTGGCCCCG GTCGTGTGGA ACCGCAACGC  
 20701 GCTCCACGAG AACGGCAGCG GCACCTCCGC TTCTGTCTCC GCGAGCAGCG GCAGGACGGT  
 20761 GACGTGCAAG GCCGCGTCTGA ACAGCGCCGG TTGGACGCCA TAGTGCGGCG TGTCTCCG  
 20821 CTGTTCCCGG GCGATCTCCA CCTCGGCGTA CAGGGTTTCG CCGTCCGCGC AGCGGTGCG  
 45 20881 CAGTCCCTGG AACGCTGGGC CGTAGCTGTA GCCGGTCTCG GCCAGCCGCT CGTAGAACGC  
 20941 GCTCACGTCC ACGCGTCCGC CGCCCGGCGG CGGCCACGCG GCGGCGGGGA CCGCCGCGAC  
 21001 GCTTCCGGCC CGGCCGAGGG TGCCGCTGGC GTGCCGGGTC CAGCTGTCCG TGCCCTCGGT  
 21061 ACGCGGCTGG ACGGTCACTC GCCCGCGTCC GGCCTCATCG GCGCTTCTGA CCGTCAACGA  
 50 21121 CACATCCACC GCGCCGCTCA CCGGCACCA GAGCGGGGTC TCGATGACCA GTTCATCCAC  
 21181 CACCCCGCGA CCGGTCTCGT CACCGGCCCG GATGACCAGC TCCACAAACG CCGTACCCGG  
 21241 CAGCAGAACC GTGCCCCGCA CCGCGTGATC AGCCAGCCAG GGATGCGTAC GCAACGAGAT  
 21301 CCGGCCASTG AGAACAACAC CACCACCGTC GTCGGCGGGC AGTGCTGTGA CCGCGGCCAG  
 21361 CATCGGATGC GCGGCCCCCG TCAGCCCGGC CGCGGACAGA TCGGTGGCAC CCGCCCGCTC  
 21421 CAGCCAGTAC CGCCTGTGCT CGAACGCGTA GGTGGGCGA TCGAGCAGCC GTCCCGGCAC  
 55 21481 CGGTCGAGCC ACGGTGTCCC AGTCCACTGC CGTGCCAGG GTCCACGCTT GCGCCACGCG  
 21541 CGTCCGCGAG CGTCCGAGC CGCCGTCACC GGTCCGCAAC GACGCCACCG TGTGAGCTG  
 21601 TTCCATCGCC GCGAGCAGCA CCGGATGGGC GCTGCACTCC ACGAACCGG ACCCGTCCAG  
 21661 CTCCGCCACC GCGCGCTCCA GCGCGACGGG GCGACGCGAG TTCCGCTACC AGTAGCCCTC  
 21721 ATCCACCGGC TCGGTCACCC AGGCGCTGTC CACCGTGGAC CACCGAGCCA CCGACCCGGT  
 60 21781 CCGCGCGGAA ATCCCTCCA GTACCTCGGC CAACTCGTCC TCGATGGCTT CCACGTGGGG

21841 CGTGTGGGAG GCGTAGTCGA CCGCGATAAG GCGCACTCGC ACGCCTTCGG CCTCGTACCG  
 21901 CGTCACCACT TCTTCCACCG CGGACGGGTC CCCC GCCACC ACAGTCGAAG ACGGGCCGTT  
 21961 ACGCGCCGCG ATCCACACGC CCTCGACCAG GTCCACCTCA CCGGCCGGCA AGGCCACCA  
 22021 AGCCATCGCC CCGCGCCCGG CCAGCCGCCC GGCGATCACC TGGCTGCGCA AGGCCACCA  
 5 22081 GCGGGCGGCG TCCTCAAGGC TGAGGGCTCC GGCCACACAC GCGCGCGCGA TCTCGCCCTG  
 22141 GGAGTGTCCG ACCACCGCGT CCGGCACGAC CCCATGCGCC TGCCACAGCG CGGCCAGGCT  
 22201 CACCGCGACC GCCCAGCTGG CCGGCTGGAC CACCTCCACC CGCTCCGCCA CATCCGGCCG  
 22261 CGCCAACATC TCCCGCACAT CCCAGCCCGT GTGCGGCAAC AACGCCC GCG CACACTCCTC  
 22321 CATACGAGCC GCGAACACCG CATCAACTCC ACACCCATGC CCACCCACTG  
 10 22381 AGCACCCCTGC CCGGGAAAGA CGAACACCGT ACGCGGCTGA TCCACGCCCA CACCCATCAC  
 22441 CCGGGCATCG CCAACAACA CCGCACGGTG ACCGAAGACA GCACGCTCAC GCACCAACCC  
 22501 CTGCGCGACC GCGGCCACAT CCACACCACC CCGCGCAGA TACCCCTCCA GCGGCTCCAC  
 22561 CTGCCCCCGC AGACTCACCT CACTCCGAGC CGACACCGGC AACGGCACCA ACCCATCGAC  
 22621 AGCCGACTCC CCACGCGACG GCCCGGGAAC ACCCTCAAGG ATCACGTGCG CGTTCGTACC  
 15 22681 GCTCACCCCG AAAGCGGAGA CACCGGCCCG GCGCGGACGT CCGCGCTCGG GCCACGCCCG  
 22741 CGCTCGGTG AGCAGTTCCA CCGCGCCCTC GTTCCAGTCC ACATGCGACG ACGGCTCGTC  
 22801 CACATGCAGC GTCTTCGGCG CGATGCCATA CCGCATCGCC ATGACCATCT TGATGACACC  
 22861 GCGACACCCC GCAGCCGCCT GCGCATGACC GATGTTGAC TCAACGAAC CCAGCAGCAG  
 22921 CGGAACCTCA CGCTCCTGCC CGTACGTGCG CAGAATCGCG TGCGCCTCGA TGGGATCGCC  
 20 22981 CAGCGTCGTC CCGTCCCGT GCGCCTCCAC CACGTCCACG TCGGCGGGGG CGAGCCCCGC  
 23041 CTTGTGGAGG GCCTGGCGGA TGACGCGCTG CTGGGAGGGG CCGTTGGGTG CGGAGATGCC  
 23101 GTTGGAGGCG CCGTCTGGT TGACGGCGGA GGAGCGGACG ACCGCGAGGA CCGTGTGTCC  
 23161 GTTGCCTCG GCGTCGGAGA GCTTTTCGAC GACGAGGACG CCGGCCCTC CGCGAAACC  
 23221 GGTGCCGTCC GCGCGCTCAG CGAACGCCCT GCACCGTCCG TCCGCGCGCA CGCCGCCCTG  
 25 23281 CCGGGAGAAC TCCACGAAGG TCTGTGGTGA TGCCATCACT GTGACACCAC CGACCAGCGC  
 23341 CAGCGAGCAC TCCCGGTCC GCAGCGCCTG CCGGCGCTGG TGCAGCGCGA CCAGCGACGA  
 23401 CGAACACGCC GTGTGACCG TGACCGCCGG ACCCTCCATG CCGAAGAAGT ACGACAGCCG  
 23461 TCCGGCGAGC ACCGCGGGCT GTGTGCTGTA GCGCGCGAAT CCGCCAGGT CCGCGCCCGT  
 23521 GCGTAGCCG TAGTAGAAGC CGCCGACGAA GACGCGGTCG TCGTGCCTG GCAGGGTGTG  
 30 23581 CCGCACGATG CCGCGTGTG CCGCGCCTC CCAGGCGATT TCGAGGAGGA TCCGCTGTG  
 23641 CCGGTGAGT GCGGTGGCCT CGCGCGGACT GATGCCGAAG AACCGCGCAT CGAAGTCGCG  
 23701 GCGCGCCGCG AGTGCGCCGG CCGCGCCCGT GCGGACTCG GCGGCGGCGT GCAGCGCGGC  
 23761 CACGTCCAG CCGCGGTGCG TGGGGAAGTC GCGGATCGCG TCGCGGCCGT CCGCGACGAG  
 23821 CTGCCACAGC TCTTCCGGTG AGGTGACGCG CCGCGGCGAGT CCGCAGGCCA TGCCGACGAC  
 35 23881 GCGAGCGGC TCGTTCGCGG CCGCGCGCAG CCGGCTGTT TCCCGCGCGA GCTGCGCGTT  
 23941 GTCCTTGACC GACGTCCGCA GCGCCTCGAT CAGGTGCTTC TCGGCCATCG CCTCATCCCT  
 24001 TCAGCACGTG CCGCATGAGC GCGTCTGCGT CCATGTCTG TAACAGTTG TCGTCCGCT  
 24061 CCGCGGTCTG GGTGCTCGCG GGTGCTGTG CCGGTGGTTC ACCGCGTCC GGGGTCCCGT  
 40 24121 TGTCGTCCGG GGTCCCGTTG ACGTCCGGGG CCAGGAGGGT CAGCAGATGA CCGGTGAGCG  
 24181 CGCCGGCGGC GGGATAGTCG AAGACGAGCG TGGCCGGCAG CGGAATGCCG AGGGCCTCGG  
 24241 AGAGCCGGTT GCGCAGGCCG AGCGCGGTGA GCGAGTCGAC CCGAGGTCC TTGAACGCCG  
 24301 TGGTGGCCGT GACCGCCGCC GCGTCCGTGT GCGCCAGCAG GGTGGCGGCG GTGTGCGGGA  
 24361 CGACGCCGAG CAGCACCTGT TCCCGTTCT TGTGGGGCAG GTCCGCGAGG CTTCCAGCA  
 45 24421 GGGAGCCGCC GTCGGTCCGCG GAGCGCCGGG TGGGGCGCTG GATCGGTGCG CACAGCGGTG  
 24481 ACGGTCGCC GGGCCCGGGT GGGGCGGTG CCACGACCAC GGCTTCCCCG GTGGCGCACG  
 24541 CCGCGTCGAG GAGGTCCGTC AGCCGGTCCG CCGCGCGGCT GAACGCCACG GCCGGCAGGC  
 24601 CTTGTGCCCC GCGCAGGTG GCGAGGGCCT GGAGCGGTCC GGCCGCTCG CCGGACGGAA  
 24661 CCGCGAGAAC GAACGCGGTC AGGTGAGGT CCGGGTTCAG GCGGTGACGT TCCAGGCCG  
 50 24721 ACTCGGCGGT CCGTCCCGG TGGACGACC CGGTACCGG GGTTCGCGG ACTGTGCCCG  
 24781 GTCGTACCG GATCACTTCG GCGCGGTGTC CCGCGAGGTG TCCGGCGAGT TCTCCGAAC  
 24841 CGCCCGCGAG GAGGACGGTG TCGCCGTAC AGGCGCGCGC CGTGGTGGGC GCGGCGGGGA  
 24901 CGAGGCGGGG CGCTTCGAGG CGCCCGTCCG CCAGGCGCAG GTGCGGTTG TCGAGGCGGG  
 24961 AGAGGGCGGC CCGCGCGCGG GGGGTGACCG TGTCGCTGCT CTCACGAGC ACGAGCCGCG  
 55 25021 CCGGTTCCCG GGTGTCGAGC AGTGCGGCGA CCGCACCGGC GACGGGCCCG GCCTCGGCGG  
 25081 ACACCACCAG CGTGGCGCCG GCGGTCTCTG GGTGCTCCAG TGCGGTACGG ACCTCGTCCG  
 25141 GACCGGATAC CCGGACGACG ATGACGTCCG CCGTGGCGTC GTCCCGAGG TCGGCTGACC  
 25201 GCGGGGCGGT GGTCCCGGGT CCGCGCGGCG CCGGACGCG GGTCCGAGG GCGGTGAACA  
 25261 CCGGCACGTC CCGTCCCGG CCGGTCTGCG CCGGGGGCGG GGTGATGAGC GAGCCATCT  
 25321 GAGCCACCGG CCGTCCAGT TCGTCCGCGA GGTGCACGCG GCGCGCCGCC TCGCCCTCGC  
 60 25381 CGTGACGAA GGTGACGCG AGTTTCGTGG CCGCGTGGT GTGGACACGG ACGCCGGTGA

SUBSTITUTE SHEET (RULE 26)

	25441	ACCGGAACGG	CAACCGTACC	CCCGCGTTCT	CGGCGGCGCG	GCCGATGCTG	CCCGCTTGCA
	25501	CGCGCGTGAC	GAGCAGCGCC	GGGTGCAGTG	TGTAGCGGGC	GGCGTCCCTG	GCGAGGGCGC
	25561	CGTCGAGGCG	GACTTCGGCG	CAGACGGTGT	CTCCGTGGCT	CCACGCGGCG	GACATGCCCG
5	25621	GGAACTCGGG	GGCGAACTCG	TATCCCGCGT	CGTCGAGTCG	CTGGTAGAAG	GCCGCGACCT
	25681	CGACCGGTTG	CGCGTGCTCG	GGCGGCCAGG	GCCCCGGCGT	GGTGGCCGGT	TCGGTGGTGG
	25741	CGATGCCGGC	GAAGCCGGAG	GCGTGCGGGG	TCCATGTCCG	GTCGCCGTCC	GTCCGGGCGT
	25801	GGACGCGCAC	GGCACGGCGT	CCGGTGTCTG	CGGGCGCGGC	GACGGTCAAC	CGCACCTGGA
	25861	CGGCGCCGGT	GGCGGGCAGG	ACCAGCGGTG	TCTCGACGAC	CAGTTCGTCT	AGCAGGTCCG
	25921	AGCCTGCCTC	GTCGGCGCCG	CGTCCGGCCA	ATTCCAGGAA	GGCGGGTCCG	GGCAGCAGTA
10	25981	CGCGCCGTC	GACGGAGTGA	CCGGCCAGCC	ATGGGTGGGT	GGCCAGCGAG	AACCGGCCGG
	26041	TGAGCAGCAC	CTCGTCCGAG	TCGGGGAGCG	CCACCGACGC	GGCGAGCAGC	GGGTGGTCCA
	26101	CGCGCTCGAG	TCCGAGGCCG	GAAGCGTCCG	TGCCGGCCGC	GGTCTCGATC	CAGTAGCGCT
	26161	CATGGTGGAA	GGCGTATGTG	GGCAGGTCTG	GTGCCGTCTG	CGTCCGCGGG	ACGACCGCCG
	26221	CCAGTCCGAC	GGGCACGCCG	GTTGTGTGCG	CCTCGGCCAG	CGCGGTGAGC	AGCCGGTGGG
15	26281	CTCCCCCGCC	GCGGCGGAGC	GTGGCGACGG	TCCGCGCGTC	GATCGCGGGG	AGCAGCACCG
	26341	CGTGCGCGCT	GACCTCGACG	AACACGGTGT	CACCCGGCTC	GCGGGCAGCG	GTACGCGCCG
	26401	TGCGGAAGCC	TACGGGGTGG	CGCATGTTGC	GGAACCAAGT	CTCGTCTGCG	AGCGGCGCGT
	26461	CGATCCAGCG	TTCGTCCGCG	GTGGAGAACC	ACGGGATCTC	GGGCGTGCGC	GAGGTGGTGT
	26521	CCCGGACGAT	CCGCTGGAGT	TCGTCTGACA	GCGGGTCGAC	GAACGGGGTG	TGGGTCCGGC
20	26581	AGTCGACGGC	GATGCGGGCG	ACCCAGACGC	CGCGGGCCTC	GTAGTCGGCG	ATCAGCGTTT
	26641	CGACGCGCTC	CGGGCGCCCG	GCGACGGTCG	TGGTGGTGGC	GCCGTTGCGG	CCCGCGACCC
	26701	AGACGCGCTC	GATCCGGGCG	GCGACGCGCT	GCGCGCGGCG	GGCCGGGAGC	CGCACCGAGC
	26761	CCATCGCGCC	GCGTCCGGCG	AGTTCGCGCA	GGAGCAGGAG	AACGCTGCGC	AGCGCAGCA
	26821	GGCGGGCACC	GTCTCCAGG	GTGAGCGCTC	CGGCGACACA	GGCCGCGGCG	ATCTCGCCCT
25	26881	GGGAGTGTCC	GATGACGGCG	TCCGGGCGTA	CGCCCGCGGC	CTCCACACAG	GCGGCCAGCG
	26941	ACACCATGAC	GGCCAGCAG	ACGGGGTGCA	CGACGTCGAC	GCGGCGGGTC	ACCTCCGGGT
	27001	CGTCGAGCAT	GGCGATGGGG	TCCCAGCCCC	TGTGCGGGAT	CAGCGCGTCT	GCGCATTGGC
	27061	GGTCTCTGGC	GGCGAACACC	GGGGAGGCCG	CCATCAGTTC	GACGCCCATG	CCGCGCCACT
30	27121	GCGTCCCTTG	TCCGGGGAAG	ACGAAGACGG	TGCGCGGCTC	GGTGAGCGCC	GTGCCCGTGA
	27181	CGACGTCGTC	GTCGAGCAGC	ACGGCGCGGT	GCGGGAACGT	CGTACGCGTC	GCGAGCAGGC
	27241	CCCGCGCGAT	GGCGCGCGGG	TCGTGGCCCG	GACGGGCGGC	GAGGTGCTCG	CGGAGTCGGC
	27301	GGACCTGGCC	GTCGAGGGCC	GTGGCGGTCC	GCGCCGAGAC	GGGCACTGGT	GTGAGCGGCG
	27361	TGCGGATCAG	CGGCTCACCG	GGCTTCGAGG	CCGACGGCTC	CTCGGCGGGC	GGCTCCCCGG
35	27421	CCGGGTGGGC	TTCCAGCAGG	ACGTGGGCGT	TGGTGCCGCT	GACGCCGAAG	GAGGACACAC
	27481	CGCGCGCGCG	CGGGCGGTCT	GTCTCGGGCC	AGGGCCGGGC	ATCGGTGAGG	AGTTCGACCG
	27541	CGCGCGCGCT	CCAGTCGACG	TGCGAGGACG	CGGTGTCCAC	GTGACGGGTG	CGCGGCAGGG
	27601	TGCGCTGCCG	CATGGCGAGG	ACCATCTTGA	TGACACCGGC	GACACCCGCG	GCGGCTGAG
	27661	TGTGGCCGAT	GTTGGACTTC	AGCGAGCCCA	GCAGCACCGG	GGTGTGCGCG	CCCTGCCCGT
40	27721	AGGTGGCCAG	CACCGCCTGT	GCCTCGATGG	GATCGCCGAG	CCTGGTGCCG	GTGCCGTGCG
	27781	CCTCCACGGC	GTCCACGTCC	GCCGGGGTGA	GCCCGGCGGT	GGCCAGGGCC	TGCCGGATCA
	27841	CCCGCTCCTG	CGAGGGCCCC	TTCCGGCGCC	ACAACCCGTT	GGAAGCACCG	TCCTGGTTGA
	27901	CCGCCGAACC	CCGGACAACC	GCCAGCACAC	GGTGGCCGTT	GCGCTCGGCA	TCGGAGAGCC
	27961	TCTCGACGAT	CAGCACACCG	GACCCCTCGG	CGAAACCGGT	GCCGTACGCC	GCATCCGCGA
45	28021	ACGCCTTGCA	GCGCGCGTCG	GGCGCGAGAC	CCCGCTGCTG	GGAGAACTCG	ACGAAGCCGG
	28081	ACGGCGAGGC	CATCACCGTG	ACGCCGCCGA	CCAGGGCGAG	CGAGCATTCG	CCGGAGCGCA
	28141	GTGACTGCCC	GGCCTGGTGC	AGCGCCACCA	GCGACGACGA	ACACGCCGTG	TCGACCGTGA
	28201	CCGCCGGACC	CTCCAGACCG	TAGAAGTACG	ACAGCCGACC	GGACAGCACA	CTGGTCTGGG
	28261	TGCCGCTCGC	GCCGAAACCG	CCCAGGTCCG	TGCCGAGTCC	GTACCCGTCC	GAGAAGGCGC
50	28321	CCATGAACAC	GCCGGTGTCT	CTTCCGCGCA	GCGACTCCGG	GAGGATCCCG	GCGTGTTCAC
	28381	GCGCCTCCCA	CGAGGTCTCC	AGGACCAGAC	GCTGCTGCGG	GTCCATCGCC	AGCGCCTCAC
	28441	GCGGACTGAT	CCCGAAGAAC	GCGCGCTCGA	AGTCCGCCAC	CCCGGCGAGG	AAGCCACCAT
	28501	GACGCACGGT	CGACGTGCCC	GGATGATCCG	GATCGGGATC	GTACAGCCCG	TCCACGTCCC
	28561	AACCCACGGT	CGTCGGAAAC	GCGGTGATCC	CGTCACCAAC	CGACTCCAGC	AGCCGCCACA
55	28621	AGTCCCTCCG	CGACGCGACC	CCACCCGGCA	GCGGGCAGGC	CATCCCCACG	ATCGCCAAAC
	28681	GCTCGTCCCT	CCGGACGGCC	GCGGTCTGTC	TGCGGTGTCG	CGATGCCGTC	CGGCCGAGCA
	28741	GCGCGCGCGT	GAGCTTCGCC	GCGACGGCGC	GCGGCGTCCG	GAAGTCGGA	ACCGCGGTGG
	28801	GCGCGAGCGG	TACGCGCGTC	GCGTCTGCTG	AGCGGTTGCG	CAGCGGATCG	CCCATGAGCG
	28861	AGTCGACGCC	GAGTTCCTTG	AACGTGCGCG	TGCGCTCGAC	CCGTGCGGCA	CCGTCTGGCG
60	28921	CGAGTACGGC	CGCGGTGCAC	TGCCGGACGA	CGGCGAGCAC	GTCTTTTTCG	GCGTCCGCGG
	28981	CGGAGAGCCG	GCGGATCCGG	TGCGCGAGGG	TGGTGGGCGC	GGCCGCCCGG	CGCCCGCGCT

SUBSTITUTE SHEET (RULE 26)

29041 CCCGGGCGCGG TGC GCGCAGC AGGGGCGAGC TGGCGAGGCC GGCCGGGTGCG GCGGCGACCA  
29101 GCGCCGGGTGCG CAGGACCGC AACGCCGCGT CGAACAGCGT CAGTCCGCCCT TCGGCGGTGCA  
29161 GCGCCGTACAC GCGGTGCGCG CGCATGCGGG CGCCGGTGGC GACCGTCAGC CGGCTCTCCG  
29221 GTTCCACACAG GCGCCAGGCC ACGGACAACG CGGGCASTCC GGCTGCCCGG CGCTGTTTGG  
5 29281 CCAGCGCSTC GAGGAACCGG TTCGCGGCCG CGTAGTTGCC CTGTCCGGGG CTSCCGAGCA  
29341 CACCGGCGGCG CGACGAGTAG AGGACGAACG CGGCCAGTTC CGTGTCTTGG GTGAGTTGGT  
29401 GCAGGTGCCA CGCGGCGTCC ACCTTCGGGC GCAGCACCGT CTCGAGCCCG TCGGGGGTGA  
29461 GCGCGGTGAG GACGCCGTCG TCGAGGACGG CGCGGGTGTG CACGACGGCC GTGAGCGGGT  
10 29521 GCGCCGGGTGCG GATCCCCGCC AGTACGGAGG CGAGTTCTTC CCGGTGCGCG AGCTCCAGG  
29581 CGATCGCGCT GACCTCGGCG CCGGGCACGT CGCTCGCGT GCGGTGCGCG GACAGCATCA  
29641 GCAGCCGGCG CACGCCGTGG CGTTCGACGA GGTGGCGGCT GATGATGCCG GCCAGCGTCC  
29701 CGGAGCCACC GGTGACGAGC ACGGTGCCGT CCGGTGCGAG CGCCGGAGCG TCACCCCGCG  
29761 GGACCGCCCG GGCCAGACGG CCGGGCGTACA CCTGGCGCTC ACGCAGCACC ACCTGGGGCT  
15 29821 CATCGAGCGC GGTGGCGGCT GCGAGCAGCG GCTCGCGGCT GTCCGGGGCG GCGTCCAGCA  
29881 GACCGATCCG GCGGGGGTGT TCGGCCGTGG CCGTCCCGAC CAGTCCGGCG GCGCGGCGCG  
29941 ACGCGAGACC GGGCCCGGTG TGGACGGCCA GGACCGCGTC GCGGTACCGG TCGTCCGTGA  
30001 GGAAGCGCTG CACGGCGGTC AGGACGCCGG CGCCCASTTC GCGGGTGTG TCGAGCGGG  
30061 CACCGCCCGC GCGGTGCGCG GGGAGGATCA CCACGTCCGG GACCGTCGGG TCGTCCAGGC  
20 30121 GGCCGGTCTG CCGGTGCGTG GCGGGCAGCT CCGGGAGCTC GGCCAGCACC GGGCGCAGCA  
30181 GGCCCGGAAC GGCTCCCGTG ATCGTCAGGG GGCGCCTGCG CACGGCGCCG ATGGTGGCGA  
30241 CCGGCCCGCG GGTCTCGTCC GCGAGGTGTA CGCCGTGAGC GGTGACGGCG ACGCGTACCG  
30301 CCGTGGCGCG GGTGGCGTGG ACGCGGACGT CGTCGAACGC GTACGGAAGG TGGTCCCGCT  
30361 CCGCGGCGAG GCGGAGTGCG GCGCGGAGCA GCGCGGGTG CAGGCCGTAC CGTCCGGCGT  
25 30421 CCGCGAGCTG TCCGTGCGCG AGGGCCACTT CCGCCAGAC GCGCTCGTCC TCGGCGCGA  
30481 CCGCGCGCGG GCGGGGCGAG GCGGGCCCGT CCGTGTACCC GGCTCGGGCC AGACGGTCCG  
30541 CGATGTCTTC GGGGTCCACC GCGCGGGCGG TGGCGGGCGG CCACGTGAGC GGCATCTCC  
30601 GCACGGCCCG GCGCGTCCCG GGTGCGGGG CGAGGATTC GTGCGCGTGC TCGGTCCACT  
30661 CCGCGCGCGC GTGCGCGCTG TGCACGGTGA CCGCGCGCGG GCGGTCCGCC CCGGGCGCGC  
30721 TCACCGTGAC CGAGAGCGCG AGCGACCGG ACCGCGGCG CGTGAGGGG GTGTCCACCG  
30 30781 TGAACGTGTC GAGGGCGCGG CAGCCGCGTT CGTCCCGCG CCGGTCCGCC AGATCCAGGA  
30841 GGGCCGCGCG GGGCAGCACC GCGAGGCGGT GCAGGGAGTG CGCCAGCGGA TCGGCGCGGT  
30901 CGACCCGGCG GGTGAGCACC AGGTGCGCGG TGCCGCGCAG GGTGACCGCC GCGGTGAGCG  
30961 CCGGGTGGCG GACCGGCGTC TGTCCGGCGG GGGCCCGCTC GCGCGCGGTG TGGGTGCCGA  
31021 GCCAGTAGCG GACCCGCTCG AACGGGTACG TCGGCGGGTG CGAGGCGCGT GCGGCGCGG  
35 31081 GGTGATGAC CTTCCGCCAG TCGACCGTGA CGCCGTGCGT GTGCAGCCCG GCGAGCGCGG  
31141 TCAGGGCGGA TCGCGTTCG TCGTCCGCGT GCAGCATCGG GATGCCGTG ACGAGTCCGG  
31201 TCAGGCTCCG GTCCGGGCGG ATCTCCAGGA GCACCGCCCC GTCGTGCGCG GCGACCTTGT  
31261 CCGCGAACCG GACGGTGTGCG CGGACCTGTC GTACCCAGTA CTCCGGCGTG GTGCAGGCGG  
40 31321 CCGCGCGCGC CATCGGGATC CTCGGCTCGT GGTACGTGAG GCTCTCCGCG ACCTTGCGGA  
31381 ACTCCTCGAG CATCGGCTCC ATCCGCGCGG AGTGAACGC GTGGCTGGTC CCGAGCGCGG  
31441 TGAAGCGGCG GAGCCGGGCG GCGACGTGCA GCACCGCTC CTCGTACCGG GAGAGCACGA  
31501 TCGACGCGGG CCGGTTGACC GCGGCGATCT CCACGCCGTC CCGCAGCAGC GGCAGCGCGT  
31561 CCGGTTCCGA CGCGATCAG GCGGCCATCG CCGCGCGGA CCGCAGCGCC TGCATCAGGC  
45 31621 GGGCCCGTGC GGACACCAGC CTGCACGCGT CCTCCAGGGA CCAGACGCGC GCGACGTACG  
31681 CCGCGGCCAG CTCGCCGATC GAATGGCCCA CGAAGGCGTC CCGGCGTACG CCGACGCGT  
31741 CGAGCTGTGC GCCGAGTGCG ACCTGGAGCG CGAACACCGC GGGCTGGGCG TACCGGTGT  
31801 CGTGAGGTC GAGCCCGCGG GGCACGTGCA GGGCGTCCAG CACCTCGCGG CGAGTGC3GG  
31861 CGAAGACGTC GTAGGCGGCG GCCAGTCCGT CGCCCAT3CC GGGACGTTGT GAGCCCTGTC  
50 31921 CGGAGAAGAG CCACACGAGG CCGCGGTCCG GTTCTCGCGC GCGGTGAGC GTGTCCGTG  
31981 CGATCAGCGC GGCCCGGTGC GGAAGGCGG TCGGGGCGAG CAGGCGCGG GCCACGCGC  
32041 GCTCGTCTC CTCGCCGGTG GCGAGGTGGG CGCGCAGGCG GTGTACCTGT GCGTCCAGTG  
32101 CCTGCGGGGT GCGTCCCGAG AGCAGCAGGG GCAGCGGTCC GGTGTGCGGT GCGGGGCGG  
32161 GTTCCGGGGC CGGTCCGGGG TGGCTTTTCA GGATGATGTG AGCTTTGGTG CCGCTAACGC  
55 32221 CGAAGGAGGA CACCCCGCGG CGCCGTGGGG GGTGCGTTTC GGGCCAGGGG CCGGCGTGG  
32281 TGAGGAGTTC GACGGCGCGG GCGGTCCAGT CGACGTGCGA GGACGGCGTG TCCACGTGCA  
32341 GGGTCCGCGG CAGGGTGCGG TGGCGATGG CGAGGACCAT CTTGATGACA CCGGCGAGCG  
32401 CCGCGCGCGG CTGAGTGTGG CCGATTTGG ACTTCAGCA GCCACGAGC ACCGGGTGT  
32461 CCGGATGCTG CCGGTAGGTG GCCAGTACCG CCTGCGGCTC GATGGGGTGC CCGGCGTGG  
60 32521 TCCCGGTGCC ATGCGCCTCG ACAGCGTCCA CATCCGCGCG GGTGAGCCCG GCGTTGGCCA  
32581 GCGCCTGCCG GATCACCCGC TCCTGCGACG GCGCGTTCCG CCGCGACAAC CCGTTGGAAG

SUBSTITUTE SHEET (RULE 26)

32641 CACCGTCTTG GTTGACCGCC GAACCACGCA CGACCGCCAG GACATTGTGG CCGTGCCGCT  
32701 CGGCGTCCGA GAGCCTCTCG ACGATCAGCA CACCGGATCC CTCGGCGAAA CCGGTGCCAT  
32761 CAGCCGCATC CGCGAACGCC TTGCAGCGGC CGTCCGGGGA GAGGCCCGCG TGCTGGGAGA  
5 32821 AGTCCACGAA GCCGGACGGC GAGGCCATCA CCGTGACGCG GCGGACCAAG GCGAGCGAGC  
32981 ACTCCCGCGA GCGCAGCGAC TGCCCGGCGT GGTGCAGCGC CACCAGCGAC GACGAACAGC  
32941 CCGTGTCCAC CGTGACCGCC GGACCCTCCA AACCCTAGAA GTACGACAGC CGACCGGACA  
33001 GCACACTGGT CTGGGTGCTG GTGGCACCGA AACCGCCGCG CTCGGCTCCA GTGCCGTACC  
33061 CGTAGAAGTA GCCGCCCATG AACACGCCGG TGTCTCTTCC GCGCAGCGAC TCCGGGAGGA  
10 33121 TCCCGGCGTG TTCCAGCGCC TCCCACGAGG TCTCCAGGAC CAGACGCTGC TGCGGGTCCA  
33181 TCGCCAGCGC CTCACGCGGA CTGATCCCGA AGAACGCCGC GTCGAAGTCC GCCACCCCGG  
33241 CGAGGAAGCC ACCATGACGC ACGGTCGACG TGCCCGGATG ATCCGGATCG GGATCGTACA  
33301 GCCCGTCCAC GTCCCAACCA CCGTCCGTCG GAAACGCCGT GATCCCGTCA CCACCCGACT  
33361 CCACGAGCGC CCACAAGTCC TCCGGCGAGC CGACCCACCC CGGCAGCCGG CAGGCCATCC  
33421 CCACGATCGC CAACGGCTCG TCCGCGCGA CGGCGCGGCT GCGGGTACGC CGCGGGTGG  
15 33481 TGGCCCGCGC GCCGGCCAGT TCGTCCAGGT GGGCGGCGAG CGCCTGCGCC GTGGGGTGGT  
33541 CGAAGACGAG CGTAGCGGGC AGCGTCAGGC CCGTCGCGTC GGCCAGCCGG TTGCGCAGTT  
33601 CGACGCGCGT CAGCGAGTCG AAGCCCACTT CCCTGAACGC GCGCGCGGGT GCGATGGCGT  
33661 GGGCGTCCGC GTGGCCGAGC ACCGCGGCGC CGCTGGTACG GACGAGGTCG AGCATGTCGC  
20 33721 GCGCGGCCGG AGGTGCGGAC GTGCGCCGGA CGGCGGCGAC GAGGGTGCCT AGGACCGGCG  
33781 GGACCGGTC GGACCGGCGC ACGCGGCGCA GGTGAGCCG GATCGGCACG AGCGGGGCC  
33841 GGTGCGGTGTG CAGGCGCGCG TCGAACAGGG CGAGCCCTG TGCGCCGCTC ATCCGGGTCA  
33901 TGCCGTTGCG GCGGATGCGG GCCAGGTCGG TGCGSGTCAG CCGCCCGCCC ATCCCGTCCG  
33961 CCGCGTCCCA CAGTCCCCAG GCGAGCGAGA CGGCGGGCAG CCCCTGGTGG TGCCGGTGGC  
25 34021 GGGCGAGCGC GTCGAGGAAC GCGTTGCCGG TCGCGTAGTT GGCCTGACCC GCGCCGCCGA  
34081 ACGTGGCGGA TATGGACGAG TACAGGACGA ACGCGGCCAG GTCGAGATCG CGCGTCAGCT  
34141 CGTGCGAGTG CCAGGCGAGC TCCGCCTTGA CCCGCGACAC GCGCTCCAC TGCTCCGGCC  
34201 GCATGGTCGT CACGCGCGCG TCGTCGACGA TCCCGGCCAT GTGCACGAGC GCGCGCAGCC  
34261 GCTGGGCGAC GTCGGCGACG ACTGCGGCGA GCTCGTCGCG GTCGACGAGC TCGGCGGCCA  
30 34321 CGTACCGCAC GCGGTGCTCC TCCGGCGTGT CGCCGGGCCG GCCGTGCGG GACACCACGA  
34381 CGACCTCGGC GGCCTCGTGC ACGGTGAGCA GGTGGTCCAC GAGGAGGCGG CCGAGCCCGC  
34441 CCGTGCCGCC GGTGACGAGG ACGGTCCCGC CCGTCAAGCG GGAGGTTCCG GTGGCCCGCG  
34501 CGACACGGCG CAGACGGGCC GCACGCGCTG TGCCGTGCGC GACCCCGAGC TGCGGCTCGT  
34561 CGCCGCGCGC GAGCCCGGCC GCTATGGCGG CCGGCGTGAT CTCGTCCGCT TCGATCAGGG  
34621 CGACGCGGCC GGGATGCTCC GTCTCCGCGG TCCGACCAG GCGCCAGCGC GCTTCTGCG  
35 34681 CCGGATCGCC GGTACGGGTG GCCACGATGA GCGGGGATCG CGCCCGGCGC GCGTCCGCGA  
34741 GCCAGGTCTG CACGGTGGTG AGCAGGTCGC GGCCAGCTC CCGGGTCCGG GCGCCGGGCG  
34801 AGGTGCCCCG GTCCCGGGT TCCACGGCCA GGACCACGAC CCGGGGGTGC TCGCGTCCG  
34861 GCACGTCGCG GAGGTACGTC CAGTCGGGGA CCGGTGACGC GGGCACGGGC ACCCAGGCGA  
34921 TCTCGAACAG CGCCTCGGCA TCGGGGTCGG CGGCCCGCAC GGTCAGGCTG TCGACGTCAA  
40 34981 GGACCGGTGA GCCGTGCTCG TCCGTGGCGA CGATGCGGAC CATGTGCGGG CCGACGCGTT  
35041 CGCAGAGCAC GCGCAGCGCG GTGCGGCGC GCGCGTGGAT CCTACGCGG GACGAGAGA  
35101 ACGCCAGCCG GCGCCGCTCC GGGTCCGTGA AGACCGTCCC GAGGGCGTGC AGGCGCGCGT  
35161 CGAGCAGCAC GGGGTGCAGC CCGTACCGGG CGTCCGTGAG CTGTTGCGCG AGGCGGACCG  
45 35221 ACGCGTAGGC GCGGCCCTCC CCGTCCACA TCGCGGTCAT GGCCCGGAAC GCGGGCCCGT  
35281 ACGAGAGCGG CAGCGCGTCG TAGAAGCCGG TCAGGTGCGC CCGGTGCGCG TCGGCGGGCG  
35341 GCCAGTCCAC GGGCTCCGCC GGACCGCCAG TGTCCACGCT CAGCGCTCCG GTCGCACTGA  
35401 GCGCCAGGGG GCGCGTGCCG GTACGGCTGT GCAGACTCAC CGACCGCCGT CCGGACACCT  
35461 CCGTTCCGAC GGTGGCCTGG ATCTCCGTGT CGCCGTGCGC GTCGACCACC ACCGGCGCGA  
50 35521 CGATGGTCAG CTCCGCGATC TCCGGCGTGC CGAGCCGGGC TCCCGCTTCG GCGAGCCAGG  
35581 CCACGAGCGC CGAGCCGGGC ACGATGACCC GGCCGTCCAC CTCGTGGTCG GCGAGCCAGG  
35641 GCTGACGGCG TACCGAGACA CCGCGGTGCG CAGCGCGCCC TCGCCGTGCG GCGAGGTCCA  
35701 CCCACGAGCC GAGCAGCGGG TGGCCGGACG TTCCCGCCGG TTCCCGCTCG ATCCAGTAGC  
35761 GGTACAGCGC GAACGGGTAC GTGGGCAGCG GCACCACCCG ACGCGTCGCG AACGACCAGG  
35821 TGACGGGCGC GCGCCGGACC CAGAGCGCG CGAGCGACCG AGTGAAGCGG TCCAGTCCCG  
55 35881 CCGCGCGCTCG CCGCAGTGTG CCGGTGACGA CCCTATGCGC ATGCGCGCGG ATCGGTCTCT  
35941 CCAGTGCGGT GGTGAGCAGC GGATGCGCGC TGACCTCGAC GAACGCGCGG TATCCGCGGT  
36001 CCGCCAGGTG GCGGTCGCG GCGGCGAAC GAACGGTGG GCGCAGGTTG TCGTACAGT  
36061 AGGCGGCGTC CCGGGGCCGG TCCAGCCAGC CCTCGTCCAC GGTGAGAGAG AACGGGACCT  
60 36121 CCGGCGTGCG CCGAGTGATG CCGGCGAGAG CGTCGAGCAG CCGCGCGCGG ATCGTTTCCA  
36181 CATGCGCGGT GTGCGACCG TAGTCGACGG CGATCCGGCG GCGCGGGGGG GTGGCGGCCA

SUBSTITUTE SHEET (RULE 26)

36241 GAGGCTCCTC CACGGCGTCG GCCGCAACCG CGACAACGAT CGACGCGGGT CCGTTGACCG  
 36301 CGGCGACCTC CAGGCGCCCG GCCCACACGG CGGCGTCGAA GTCGGCGGGC GGCACCGAGA  
 36361 CCATGCCGCC CTGCCCGGCC AGTTCCGGTG CGACGAGTCG GCTGCGCACC GCGACGACCT  
 36421 TCGCGGCGTC GTCCAGGGTG AGCACCCCG CGACGAGGCG CGCGGGGACT TCGCCCTGGG  
 5 36481 AGTGGCCGAC GACCGCGGCC GGGGCGACCC CGTGCGCACG CCACAGCTCC GCCAGCGCCA  
 36541 CCATCACCGC GAACGACCGG GGCTGCACGA CATCGACCCG GTCGAACCGG GGCCTCCCG  
 36601 GCGGCTGGGC GATGACGTCC AGCAGGTCCC ATCCGGTGTG CGGGGCGAGC GCGGTGGCGC  
 36661 ACTCGCGGAG CCGCCGGGCG AACACGGGCT CCGTGCGGAG CAGTTCCGCA CCGATGCGCG  
 36721 CCGACTGGGA GCCCTGCCCG GGGAAACGCG ACACGACACG TGTGTGGGTG ACGTCCGCGG  
 10 36781 TCGCCGTCAC GGCCTCCGCG ACTTCGCGAC CACGGGCGAA CGCTCCGCG TCTCGGGCGG  
 36841 CCACGACCGC CCGGTGGCGC ATGGCCGTC GGGTGGTGGC GAGCGAGTGG CCGACCGCGG  
 36901 CCGCGGCGCC AGTGAGCGGG GCCAGCTGTC CCGCGACGTC CCGCAGTCCC TCGGGGGTCC  
 36961 CGGCGACAT CGGCCAGACC ACGTCCGCG GCACCGGCTC GGCTTCGGGT GCGGACACGG  
 37021 GTGCGGGGCG GCGGGGGGCG CCGGCTCCGA GGACGACATG GCGTGTGGTG CCGCTGATGC  
 15 37081 CGAACGACGA GACACCCGCA CGCCGGGCGC GCCCGGTGAC CGGCCACGGC TCACTGCGGT  
 37141 GCAGCAGCCG GATGTCGCGG TCCAGTCCGA CGTGCCGGGA CCGCTCGTCG ACGTGCAGCG  
 37201 CCGCGCGCAG GACGCGGTGC CGCATCGGCA TGACCATCTT GATGACGCGG GCGACGCGCG  
 37261 CCGCGGCGTG GGTGTGGCCG ATGTTGCACT TGAGCGAGCC GATCAGCAGC GATCGACGCG  
 37321 GTTCGCGCCC GTAGGCCACT TGCAGGGCCT GGGCCTCGAC GGGGTGCGCG AGACGGGTGC  
 20 37381 CGGTGCCGTG TGCTCCACG GCGTCGACGT CACCCGGCGC CAGGCCGGCG TCGGCGAGCG  
 37441 CACGCTGGAT GACGCGCTGC TCGCGAGGCC CGTTCGGGGC GGACAGCCCG TTCGACGCGC  
 37501 CGTCGGAGTT GACCGCGGAG CCGCGCACCA GCGCCAGCAC GGGGTGGCCG TGGCGGGTGG  
 37561 CGTCGGAGAG CCGCTCCAGC ACCAGGACAC CGGCGCCCTC GGCGAAGCTC GTGCCGTCCG  
 37621 CGGTGCTCCG GAAGGCCTTG GCACGCGGCT CGGGGGCGAG CCGCGCTGCG CCGGAGAACT  
 25 37681 CGACGAACCC GGTGCTCGTC GCCATCACCG TGACACCGCC GACGAGGCG AGCGAGCACT  
 37741 CCGCCGAGCG CAGCGACCGC GCGGCTGGT GCAGCGCCAC CAGCGACGAC GAACACGCGC  
 37801 TGTGACGGT GACCGACGGG CCTCCAGAC CGAAGTAGTA CGAGAGCCCG CCGGAGAGAA  
 37861 CGCTGGTCCG CGTGCCGGTC GCGCCGAAAC CGCCAGGTC CACGCCCCCG CCGTAGCCCT  
 37921 GCGTGAACGC GCCCATGAAT ACGCCGGTGT CGCTGCCGCG GACGCTTTCG GGCAGGATGC  
 30 37981 CGCTCGTTTC GAACGCCTTC CACGACGCTT CGAGGACCAG ACGTGTCTGC GGGTCCATCG  
 38041 CGAGCGCCTC ACAGCGGCTG ATCCCGAAGA ACGCGGCGTC GAAGTCGGCG GTGGCGGTGA  
 38101 GGAAGCCGCC GTGACGCACG GAAACCTTGC CGACCGCGTC GGGGTTCGGG TCGTAGAGCG  
 38161 CCGCGAGGTC CCAGCCGCGG TCGGCGGGGA ACTCGGTGAT CGCGTCCCCG CCGGAGTCGA  
 38221 CGAGCCGCCA CAGGTCCTCC GGTGACCGCA CGCCACCGGG CATCCGGCAC GCCATGGCCA  
 35 38281 CGATCGCCAG CCGCTCGTTC CCGCCACCG TCGGTGCGGG CACTGTGCGC GCGGAGCGG  
 38341 CAGGGGCGCG CTCACCCGCG CGTTCCTCAT CCAGGCGGGC GCGGAGCGCG GCGGTGTGCG  
 38401 CGTGGTCGAA GACGGCCGTC GCGGAGAGCC GTACCCCGCT CGTCTCGGG AGGCTGTTGC  
 38461 GCAACCGGAC ACCGCTGAGC ACCGCTGATC CGAGGTCTTT GAACGCGGTC GTGGCGGTGA  
 38521 TCTCGGAGGC GTCGGCGTGG CCGAGCACCG CCGCCGTGGC CGCACACACG ATGGCCAGCA  
 40 38581 GTTCACGATC GCGGTGCGCG TCGCGGTGCG GGTGTCTCTC CGCACGGGCG GCGATGCGGC  
 38641 GTTCGGTCCG CTGCCGGACG GGCTCGGTGG GAATCGCCCG GACCATGAAC GGCACGTCG  
 38701 CCGCGAGGCT CCGTTCGATG AAGTGGGTGC CCTCGGCCCT GGTGAGCGGC CGGAACCCGT  
 38761 CGCGCACCCG CTGCCGTCG GCGTCTGCAA GTTGTCCGT GAGGGTGTCT GTGGTGTGCC  
 38821 ACATGCCCCA GCGGATGGAG GTGGCGGGTT GCGGAGGGT GTGGCGGTGG GTGGCGAGGG  
 45 38881 CCGCAGGAA GCGGTGGCG GCGGCGTAGT TTCCTTGTCC GGGGTGCGG AGGACGGCGG  
 38941 CCGCGCTGGA GTAGAGGACG AAGTGGGTGA GGGGTGGTT TTGGGTGAGG TGGTGCAGGT  
 39001 GCGAGGCGGC GTTGGCTTTG GGGTGGAGGA CCGTGGTGAG GCGGTGCGGG GTGAGGGCGT  
 39061 CGAGGATGCC GTCTGCGAGG GTGGCGGCGG TGTGGAAGAC GCGGTGAGG GGTGGGGGA  
 50 39121 TGTGGGCGAG GGTGGTGGCG AGTGGTGGG GGTGCGCGAC GTGCGAGGGG AGGTGGGTGC  
 39181 CGGGGTGGT GTCGGGGGT GGGGTGCGCG AGAGGAGGTA GGTGTGGGGT TGGTTCAGGT  
 39241 GCGGGGCGAG GATGCCGCG AGGGTGGCG AGCCGCGGT GATGATGATG GCGTGTTCGG  
 39301 GTTGAGGGG GGTGGTGGTG GGTGGGTGG TGGTGTGAG GGGGTGAGG TGGGTGCGT  
 39361 GAGGGGTGT GTGGGTGAGG CCGAGGTGG GGTGTGCGAG GGTGGCGAGT TGGGCCAGGG  
 39421 GAGGGGAGT GTGGGGGTGG TCGGTTTCCA TGAGGCGGAT GCGGTGGGGT TGTTCGTTCT  
 55 39481 GCGCGGTGCG GTGAGGCGG GTGACGCTGG CCGCGGCGGG GTCGGTGGTG GTGTGACGA  
 39541 TGAGGGTGT GTCGGTGGTG GTGAGGTGTT GTTGAGGGG GTTCAGGACG CCGGTGGCGG  
 39601 CGGTGTGGG CCGGTGGGT ATGCTCTGG GGTGCTCGCG GTGGGCGGCG GTGATCAGGA  
 39661 CTGTCCCTC CGGAGGTCA CCGCTCTAGA CCGCTCGGC GACCGCGAGC CACTCGAACC  
 39721 CGAGCGGGT CCGCCCCGAC GGGGTGTGG CCGCTCCCT CAGCACCAGC GAGTCCACCG  
 60 39781 ACACGACAGG ACGGCCATCC GGGTGGGCA CCGCGACGGC GACGCGGGC TCCCCCGGG

SUBSTITUTE SHEET (RULE 26)



39841 TTAGGGGCGAC GCGCACCGCG GCGGCCCGCG TGGCGTTGAG GCGCACGCCC GTCCAGGAGA  
 39901 ACGGCAGCTC GATCCCGCCG CCCGCGTCGA GGCGCCCGGC GTGCAGGGCC GCGTCGAGCA  
 39961 GTGCCGGATG CACACCGAAA CCGTCCGCCT CGGCGGCCTG CTCGTCGGGC AGCGCCACCT  
 40021 CGGCATACAC GGTGTCACCA TCACGCGAGG CAGCCCGGAA CCCCTGGAAC GCGGACCCCT  
 5 40081 ACTCATAACC GGCATCCCGC AGTTCGTGTC AGAACCCCGA GACGTCGACG GCGCGCGGCC  
 40141 TGGCCCGGGG CCACTGCGAG AACGGCTCAC CGGAAGCGTT GGAGGTATCC GGGGTGTCGG  
 40201 GGGTCAGGGT GCCGCTGGCG TGCCGGGTCC AGCTGCCCGT GCCCTCGGTA CGCGCGTGGA  
 40261 CGGTCACCGG CCGCCGTCCG GCCTCATCGG CCCCTTCCAC GGTACCCGAC ACATCCACCG  
 10 40321 CTGCGGTGTC CGGCACCACG AGCGGGGATT CGATGACCAG TTCATCCACC ACCCGCAAC  
 40381 CGGTCTCGTC ACCGGCCCGG ATGACCAGCT CCACAAACGC CGTACCCGGC AGCAGAACCG  
 40441 TGCCCCGCAC CGCGTGATCA GCCAGCCAGG GATGCGTACG CAATGAGATC CGGCGCGTGA  
 40501 GACACAACACC ACCACCGTCG TCGGCGGGCA GTGCTGTGAC GGCGGCCAGC ATCGGATGCG  
 40561 CCGCCCGGGT CAGCCCGGGC GCGGACAGGT CGGTGGCACC GGCCGCCTCC AGCCAGTACC  
 15 40621 GCGTGTGCTC GAACGCGTAG CCGCGAGCCG CCAGCAGCCG CCGCGGACCC GGTTCGACCA  
 40681 CCGTGCCCCA GTCCACCCCG GCACCCAGAG TCCACGCTG CGCCAACGCC CCCAGCCACC  
 40741 GCTCCAGGCC ACCGTACCA GTCCGCAACG ACGCCACCGT GCGGGCCTGT TCCATCGCCG  
 40801 GCAGCAGCAC CGGATGGGCA CTGCACTCCA CGAACACCGA CCCGTCCAGC TCCGCCACCG  
 40861 CGCATCCAG CGCGACAGGG CGACGAGGT TCCGGTACCA GTACCCCTCA TCCACCGCT  
 20 40921 CGGTACCCCA GCGGCTGTCC ACGGTGAGCC ACCACGCCAC CGACCCGGTC CCGCCGAAA  
 40981 TCCCTTCAG TACCTCAGC AGTTCGTCTT CGATGGCTC CACGTGAGGC GTGTGGGAGG  
 41041 CCGTAGTCGAC CGCGATACGA CGCCACCGCA CCCATCAGC CTCATACCGC GCCACCACCT  
 41101 CCTCCACCGC CGACGGGTCC CCGCCACCA CCGTCGAAGC CGGACCATTA CGGCCGCGA  
 41161 TCCACACACC CTCGACCAGA CCCACCTCAC CGGCCGGCAA CGCCACCAGG GCCATCGCCC  
 25 41221 CCGCGCCGGC CAGCCGCGCC GCGATCACCC GACTGCGCAA CGCCACCAGC CGGGCGCGGT  
 41281 CCTCCAGGCT GAGGGCTCCG GCCACACACG CCGCCGCGAT CTCCCCTGC GAGTGTCCGA  
 41341 CCACAGCGTC CGGCACGACC CCATGCGCCT GCCACAGCGC GGCCAGGCTC ACCGCGACCG  
 41401 CCGAGCTGGC CGGCTGGACC ACCTCCACCC GCTCCGCCAC ATCCGACCGC GACAACATCT  
 41461 CCGGCACATC CCAGCCCGTG TCAGGCAACA ACGCCCGCGC ACATCTCTCC ATACGAGCCG  
 30 41521 CGAACACCGC GGAACGGTCC ATGAGTTCCA CGCCCATGCC CACCCACTGG GCACCTGCC  
 41581 CGGGGAAGAC GAACACCGTA CGCGGCTGAT CCACCGCCAC ACCCATCACC CGGGCATCAC  
 41641 CCAGCAGCAC CGCAGCGTGA CCGAAGACAG CACGCTCAGC CACCAACCCC TGCGCGACCG  
 41701 CGGCCACATC CACCCACCCC CCGCGCAGAT ACCCTCCAG CCGCTCCACC TGCCCCGCA  
 41761 GACTCACCTC ACCACGAGCC GACACCGGCA ACGGCACCAA CCCATCACC CCGGACTCCA  
 35 41821 CACCGGACGG CCCAGGAACA CCTCCAGGA TCACGTGCGC GTTCGTACCG CTCACCCCGA  
 41881 ACAGCAGCAC ACCCGCATGC GGTGCCCGAT CGGACTCGGG CCAGCTCCCT GCGCTCGTGA  
 41941 GCAGCTCCAC CGCACCGGCC GACCACTCCA CATGCGACGA CGGCTCGTCC ACGTGCAGCG  
 42001 TCTTCGGCGC GATCCCATGC CGCATCGCCA TGACCATCTT GATGACACCG GCGACACCCG  
 42061 CAGCCCGCTG CGCATGACCG ATGTTGCGACT TGACCGAACC GAGGTAGAGC GCGGTGTCGC  
 40 42121 GGTCTGCCC GTAGGCCGCG AGGACGGCCT GCGCCTCGAT CCGGTGCGCC AGCGCGTGC  
 42181 CGGTGCCGTG CGCCTCCACC ACGTCCACAT CGGCGGCGCG CAGTCCGGCG TTGACCAACG  
 42241 CCGTCCGGAT CACGCGCTGC TGGCGACGCG CGTTGGGGGC GGACAGTCCG TTGGAGGCAC  
 42301 CGTCTGGTT CACCGCCGAG CCGCGGACGA CCGCGAGAAC GGTGTGCCCG TTGCGCTCGG  
 42361 CGTCGGAGAG CCGCTCCAGC ACGAGAACGC CGACGCCCTC GGCGAAGCCG GTCCCGTCCG  
 42421 CCGCGTCGGC GAACGCCCTT CACCGTCCGT CCGGGGAGAG TCCGCGCTGC CCGGAGAACT  
 45 42481 CCACGAGCTC TGCGGTGTTT GCCATGACGG TGACACCGCC GACCAGCGCC AGGGAGCACT  
 42541 CCGCGGCCCG CAGTGCTGT GCGCCTGGT GCAGGGCGAC CAGCGACGAC GAGCACGCCG  
 42601 TGTCGACCGT GACCGCCGGG CCCTGAAGTC CGTACACGTA CGAGAGGCGC CCGGACAGGA  
 42661 CGCTCGTCTG CGTCGCCGTG ACACCGAGCC CGCCAGGTG CCGGCGGACG CCGTAGCCCT  
 50 42721 GGTGGAACGC GCGCATGAAC ACGCCGGTGT CGCTCTCCCG GAGCCTGTCC GGCACGATGC  
 42781 CGGCGTTCTC GAACGCCCTC CAGGAGGTCT CCAGGATCAG GCGCTGCTG GGGTCCATCG  
 42841 CCAGCGCCTC GTTCGGACTG ATGCCGAAGA ACGCGGCGTC GAACCCGGCG CCGGCCAGGA  
 42901 ATCCGCGCTG GCGTGTGCTG GAGCGGCCCG CCGCGTCCCG GTCCGGGTG TACAGCGCGT  
 42961 CCGCGTCCCA GCGCCGGTCC GTGGGGAAC CCGTGATCGC CTCGGTACCG GCGGCGACGA  
 55 43021 GCGGCCACAG GTCTCCCGG CAGGCGAACC CCGCGGGCAG TCGGCACGCC ATGCCGACGA  
 43081 TCGCGACGGG GTGCGCGGAG GCGGCGGTCT GCGCGGTGCG GGGTGGCGCT GCGCGGAGC  
 43141 CCGCGAGGTG GCGGCGGAAC GCACGCGGAG TCGGGTGGTC GAACGCGGTT GACGCGGGCA  
 43201 CCGCGAGACC CCGCGCGCG GCGAGGCTGT TCGTGAACCT GACGGTGGTG AGCGAGTCCA  
 43261 GCGCGTCTC GCGGAACGTC CCGTCCCGGG AGCAGTGTCC GCGCGCCGGC AGGCCAGGA  
 43321 CGGTGGCGAC GCTGTGCGCG ACCAGGTGCA GCGTACGTC CTCCCGGCC GCACGGGCCG  
 60 43381 CCGCGAGGCG GTTCGCCCC TCGTGTTCG TGGCGTCGGG CTCGGCCGGT CCGGTGAGTC

SUBSTITUTE SHEET (RULE 26)



	43441	CGGTGAGGAT	CGGCGGCGTG	GCGCCCGCCA	TGGTGGGCGC	CCGCGCCCGG	CGGGAACCGG
	43501	TCCGGGCGAC	GATGTACGAG	CCGCGCGCGG	CGATGGCCTT	CTCGATCAGG	TCCCGGGTGA
	43561	GCGCGGCGCG	TTCGATGCCG	GGCAGCGCGC	GGACGGTGAC	GGTGGGGAGT	CCCTCCGCGG
5	43621	CCCGTGGCGG	GGTGTGGGCG	TCGGCGCGCG	CCGGGCGGTC	GAGCAGGACG	TGCACGAGCG
	43681	CGCCGGGGTT	CGCGGCTTCC	TCGGCTGCGG	TGGTCACGTG	GGTGAGGCGG	GTCTCGTCGC
	43741	GGAGCAGGCC	GGCGACGGTG	TCGGCGTCTT	CCCCGGTGAC	CAGGACCGGC	GCGTCCGGGC
	43801	CGATCGGAGG	CGGCACGGTG	AGGACCATCT	TGCCGGTGTG	CCGGGCGTGG	CTCATCCACG
	43861	CGAACGCGTC	CCGCGCACGG	CGGATGTCCC	ACGGCTGCAC	CGGCAGCGGG	CACAGCTCAC
10	43921	CGCGGTGCGA	CAGGTCGAGG	ATGGGCACGG	GGATCTCCCG	CAGGCGCGCG	GGATCCACST
	43981	CGGCCAGGTC	GAACGGCTGC	TGGGCGGCGT	GGCGGATGTC	GGTCTTGCCC	ATCTCGACGA
	44041	ACCGGCCGCG	CGGTGCGAGC	AGGCCGATGG	ACGCGTCGAG	GAGTTCACCG	GTGAGCGAGT
	44101	TGAGCACGAC	GTCGACCGGC	GGGAAGGTGT	CGGCGAACGC	GGCGCTGCGG	GAGTTCGCCA
	44161	CATGGTCGCT	GTCGAAGCCG	TCGGCGTGCA	GCAGGTGTTG	TTTGGCGGGA	CTGGCGGTGG
	44221	CGTACACCTC	GGCGCCGAGG	TGGCGGGCGA	TCCGGGTGCG	CGCCATGCCG	ACACCGCCCG
15	44281	TGCGGCGGTC	GACCAAGACC	TTCTGGCCCG	GTCGCAGCTC	GCCC CGCTCG	ACGAGGCCGT
	44341	ACCAGGCGGT	GGCGAACACG	ATGGGCACGG	ACGCGGCGAT	GGGGAACGAC	CATCCCCGTG
	44401	GGATCCGTGC	GACCAAGCCG	CGGTCCGCGA	CCACGCTGCG	CCGGAACGCG	TCTGACGA
	44461	GACCGAACAC	GCGGTGCGCG	GGGGCCAGGT	CGTCGACGCC	GGGTCCGACT	TCGGTCACGA
20	44521	TGCCCCGCGG	CTCCCCGCCC	ATCTCGCCCT	CGCCCCGGTA	GGTGCCGAGC	GCGATCAGCA
	44581	CGTCGCGGAA	GTTCAGCCCC	GCGGCGCGGA	CGTCGATGCG	GACCTCGCCG	GCGGCCAGGG
	44641	GCGCGGCGGG	ACGTCGAGCG	GGGCGACGAC	GAGGTCGCGG	AGCGTTCCCG	AGGCGGCGCG
	44701	GCGCAGCGCG	CACCTGCGCG	GTCCGACGGG	GGGTGGTGTG	CGCGCGTACC	AGCCGGGGCA
	44761	CGTAGGCCAC	GCCGGCCCGC	AGCGCGATCT	GGGTTTCGCC	GAGCGAGGCG	GCGGCGGGGA
25	44821	CGAGGTCTGC	ATCGCCGTCC	GTGTCCACCA	GCACGAACGA	TCCGGGTTCG	GCGGCGTGGC
	44881	GGCGCAGCGC	CTCGTCCCAG	AGCCGGGCCT	GGTCCGCGTC	CGGGATCTCG	GCCGGGCCGA
	44941	CGCCACCCGC	GCGGCGGGTG	ACGACCGTCC	GGCGGGGTGA	CGGGGTGCCG	GCGAGGTCGC
	45001	GCCGCTCCCA	GACCAAGTTC	CACAGCGTGG	CCTCGCCACT	GCCGGTGGCG	ACCAGATGGG
	45061	CCGGCAGCCC	CGCGAGCCCG	GCGCGCTGGA	CCTTGCCCGA	CGCGGTGCGG	GGGATCGTGG
30	45121	TGACGTGCCA	GATCTCGTCG	GGCACCTTGA	AGTAGGCGAG	CCGCGCGCGG	CACCTGGCGA
	45181	GGATCGCCTC	GGCGGGGACG	CGGGGGCCGT	CGGAAACGAC	GTAGAGCAGC	GATATGTCGC
	45241	CGAGGACGGG	GTGCGGGCGG	CCCGCCGCGG	CGGCGTCCCG	GACACCGGCC	ACCTCCTGGG
	45301	CGACGGTCTC	GATCTCCCGG	GGGTGGATGT	TCTCCCCGCC	GCGGATGATC	AGCTCCTTGA
	45361	CCCGGCCCGT	GATCGTCACG	TGTCCGGTCT	CGGCCTGACG	TGCGAGGTCC	CCGGTGCGGT
35	45421	ACCAGCCGTC	CACGAGCACC	TGGGCGGTCT	CCTCCGGCTG	GGCGTGGTAG	CCGAGCATGA
	45481	GGTCTGGCCC	GCTCGCCAC	AGCTCGCCCT	CCTCGCCGGG	TGCCACGTGC	GCGCCGGACA
	45541	CCGGGTGCGC	GAACCGCAGC	GACAGGCCCG	CGACGGGACG	CGCCACGACG	CCGGGAACCC
	45601	GCGCATCCTC	CAGGGTGTTC	GCGGTGAGCG	AGCCGGTCTG	CTCGGTGACG	CCGTACGTGT
	45661	CGAGCAGGGG	CACGCCGAAC	GTGCGCTCGA	AATCCCTGGT	GAGCGACGCC	GGCGAGGTGG
40	45721	ATCCGGCGAC	CAGCGCCACG	CGCAGCGCGC	GAGCCCGCGG	CTCGCCGGAC	ACGGCGCCGA
	45781	GGAGGTAGCG	GTACATCGTC	GGCACGCCGA	CGAGCACGGT	GCTGGAGTGT	TCGGCCAGGG
	45841	CGTCGAGGAC	GTCACGCGCG	ACGAAGCCGC	CCAGGATACG	GGCGGACGCG	CCGACCGTGA
	45901	GGACGGCGAG	CAGGCAGAGG	TGGTGGCCGA	GGCTGTGGAA	CAGCGGGGCG	GGCCAGAGCA
	45961	GTTCTGTCGT	CTCGGTGAGC	CGCCAGGACG	GCACGTGCGA	GTGCATCGCG	GACCAAGGC
45	46021	CGCTGCGCTG	TGCGGAAACC	ACGCCCTTGG	GACGGCCGGT	GGTGCCGGAG	GTGTAGAGCA
	46081	TCCAGGCGGG	TTCGTCCAGG	CCGAGGTCTG	CGCGGGGCGG	GCACGGCGGC	TCGGTCCCGG
	46141	CGAGGTCTCT	GTAGGAGACG	CAGTCCGGTG	CCCGGCGCCC	GACGAGCACG	ACGGTGGCGT
	46201	CGGTGCCGGT	GCGGCGCACC	TGGTCGAGGT	GGGTTTCGTC	GGTGACCAGC	ACGGTCCGCG
	46261	CGGAGTCCGT	CAGGAAGTGG	GCGAGTTCGG	CGTCGGCGGC	GTCCGGGTTC	AGCGGGACGG
50	46321	CGACGGCGGC	GGCGCGGGCG	GCGGCGAGGT	AGACCTCGAT	GGTCTCGATC	CGGTGGCCGA
	46381	GCAGCATCGC	GACCCGGTCC	CCGCGGTGCA	CGCCGGACGC	GGCGAGGTGT	CCGGCGAGCC
	46441	GGCCGGCCCG	GAGCCGGAGT	TGCGTGTACG	TCACGGCGCG	TTGGGAATCC	GTGTAGGCGA
	46501	TCCGGTTCGC	GCGTCTCTCG	GCATGGATGC	GGAGCAATTC	GTGCAACGGC	CGGATTGGTT
	46561	CCACACGGCG	CATGGAAACA	CCTTTCTCTC	GACCAACCGC	ACAACAGCAC	GGAACCGGCG
55	46621	ACGAGTAGAC	GCCGGCGACG	CTAGCAGCGT	TTTCCGGACC	GCCACCCCTT	GAAGATCCCT
	46681	CTACCGTGCC	CGGCCTCCCC	GGACGCTCAT	CTAGGGGGTT	GCACGCATAC	CGCCGTGCGT
	46741	AATTGCTTTC	CTGATGACCG	ATGCCGAGCG	CCAGGGAAGG	GTGGAGGCTT	TCTCATATC
	46801	TGTCACGGCG	CCGTATTGCC	GCTTCGAGAA	GACCGGATCA	CCGACCTCG	AGGGTGACGA
	46861	GACGGTGCTC	GGCCTGATCG	AGCACGGCAC	CGGCCACACC	GACGTGTGCG	TGGTGGACCG
60	46921	TGCTCCCCCG	ACCGCCGTGC	ACACCACGAC	CCGTGACGAC	GAGGCGTTCA	CCGAGGTCTG
	46981	GCACGCACAG	CGCCCTGTCT	AGTCCGGCAT	GGACAACGGC	ATCGCCTGGG	CCCGCACCGA

SUBSTITUTE SHEET (RULE 26)

47041 CGCGTACCTG TTCGGTGTGCG TGCGCACCGG CGAGAGCGGC AGGTACGCGG ATGCCACCGC  
47101 GGCCCTCTAC ACGAACGTCT TCCAGCTCAC CCGGTGCGTG GGGTATCCCC TGCTCGCCCG  
47161 GACCTGGAAC TACGTCAGCG GTATCAACAC GACGAACGCG GACGGGCTGG AGGTGTACCG  
5 47221 GGACTTCTGC GTGGGCGCGG CCCAGGCGCT CGACGAGGGC GGGATCGACC CGGCCACCAT  
47281 GCGCGCGCGG ACCGGTATCG GCGCCACGCG GGGCGGCATC ACCTGCGTGT TCCTCGCCCG  
47341 CCGGGGCGGA GTGCGGATCA ACATCGAGAA CCCC GCCGTC CTCACGGCCC ACCACTACCC  
47401 GACGACGTAC GGTCCGCGGC CCCC GGTTCTT CGCACGGGCC ACCTGGCTGG GCCCGCCGGA  
47461 GGGGGGCGCG CTGTTTATCT CCGCGACGCG CGGCATCCTC GGACACCGAA CGGTGCACCA  
10 47521 CCGTGTATGTG ACCGGCCAGT GCGAGGTGCG CCTCGACAAC ATGGCCCGGG TCATCGGCGC  
47581 GGAGAACCTG CCGCGCCACG GCGTCCAGCG GGGGCACGTC CTCGCCGACG TGGACCACCT  
47641 CAAGGTCTAC GTCCGCCGCC GCGAGGATCT CGATACGGTC CGCCGGGTCT GCGCCGACG  
47701 CCTGTGAGC ACCGCGGCCG TCGCCCTTTT GCACACCGAC ATAGCCCGCG AGGATCTGCT  
47761 CGTCGAAATC GAAGGCATGG TGGCGTGACA ATACCCGGTA AAAGGCCCGC GACGCTGCGC  
15 47821 CTCGGCGGAT CCGCGAAGAG AAAGAAGAGC GTCACCGCAC AGCGCGGCAG CCCGGTCTTT  
47881 TCGTCTTCG CACAGCGCGG GATCTGGTTT CTCCAGCAAT TGGACCCGGA GAGCAACGCC  
47941 TATAATCTCC CGCTCGTGCA ACGCCTGCGC GGTCTATTGG ACGCGCCGCG CCGTGGCGT  
48001 GCGCTGCGC TCGTCTGCGC GCGCCACGAG GCGTTGCGGA CCGTGTTCGA CACCGCCGAC  
48061 GCGGAGCCCC TCCAGCGGGT GCTTCCCGCC CCGGAACACC TCCTGCGCCA CGCGCGGGCG  
20 48121 GGCAGCGAGG AGGACGCCCG CCGGCTCGTC CGCGACGAGA TCGCCGCGCC GTTCGACCTC  
48181 GCCACCGGGC CGTTGATCAG GGCCCTGCTG ATCCGCTCG GTGACGACGA CCACGTTCTC  
48241 GCGGTGACCG TGCACCATGT CGCCGCGCAC GCTTGGTCTT TCGGGCTCCT CCAACATGAA  
48301 CTCGCAGCCC ACTACACGGC GCTGCGCGAC ACTGCCCGCC CTGCCGAAC TCCGCCGTTG  
48361 CCGGTGCACT ACGCCGACTT CGCCGCTGCG GAGCGGCGCG AACTCACCAG CGCCGACTG  
25 48421 GACAGGCGTC TGGCCTACTG GCGCGAGCAA CTCCGGGGCG CCGCGGCGCG GCTCGCCCTC  
48481 CCCACCGACC GTCCCCGCCG GCGGTGCGCC GACGCGGACG CCGGCATGGC CGAGTGGCGG  
48541 CCGCCGCGCG CGCTGGCCAC CGCGGTCTCT ACGCTCGCGC GCGACTCCGG TGCGTCCGTG  
48601 TTCATGACCC TGCTGGCGCG CTTCCAAGCG GTCTCGCCC GGCAGGCGGG CACGCGGGAC  
48661 TGCTGGTTCG GCACGCCCGT GGCGAACCGT ACGCGGGCGG CGTACGAGGG CCGTGTGCGC  
30 48721 ATGTTCTGCA ACACGCTCGC GCTGCGCGGC GACCTCTCGG GCGATCCGTC GTTCGGGAA  
48781 CTCCTCGACC GCTGCCGGGC CACGACCACG GACGCGTTTCG CCCACGCCGA CCTGCCGTTT  
48841 GAGAACGTCA TCGAACTCGT CGCACCGGAA CGCGACCTGT CCGTCAACCC GGTGCTCCAG  
48901 GTGCTGTTGC AGGTGCTGCG GCGCGACGCG GCGACGGCCG CGCTGCCCGG CATCGCGGCC  
48961 GAACCGTTCC GCACCGGACG CTGGTTTACC CGCTTCGACC TCGAATTCCA TGTGTACGAG  
35 49021 GAGCCGGTG GCGCGCTGAC CGCGCACTG CTCTACAGCC GTGCGCTGTT CGACGAGCCA  
49081 CCGATCACGG GGTGCTGGA GGAGTTACG CCGGTGCTTC AGGCGGTAC CGCCGACCCG  
49141 GACGTACGGC TGTGCGGGCT GCCGGCCGGC GACGCGACGG CCGCAGCGCC CGTGGTGCCC  
49201 TCGAACGACA CGGCGCGGGA CCTGCCCGTC GACACGCTGC CCGGCTGCTT GGCCCGGTAC  
49261 GCGCGACGCA CCCCCGCGCG CGTGGCCGTC ACCGACCCGC ACATCTCCCT CACCTACGCG  
49321 CAGCTGGAAC GCGGGGCGAA CCGCCTCGCG CACCTGCTCC GCGCGCGCGG CACCGCCACC  
40 49381 GCGCACTGCG TCGGGATCTG CGCCGATGCG GCGCGCGACC TGATCGTCCG CATCGTGGGG  
49441 ATCTCTAAG CGGCGCGCGC TTATGTGCGC CTGGACCCCG AACATCTCTT GGAGCGCAGC  
49501 GCGTTCTGTC TGGCCGACGC CGATGTGACC ACGGTGGTGG CGCACGAGGT TACCGTTCC  
49561 CCGTTCCCCG ATGTGCCGCA CGTGGTGGCG TTGGACGACC CGGAGCTGGA CCGGCAGCCG  
49621 GACGACACGG CGCCGACGCT CGAGCTGGAC CCGGACAGCC TCGCCTACGC GATCTACAG  
45 49681 TCCGGGTGCA CCGGCAGGCC GAAGGCCGTG CTCATGCCGG GTGTACGCG CGTCAACCTG  
49741 CTGCTCTGGC AGGAGCGCAC GATGGGCCGC GAGCCGGCCA GCCGCACCGT CCAGTTCTGT  
49801 ACGCCACGCT TCGACTACTT GGTGCAGGAG ATCTTTTCCG CGTGTCTGGG CGGCACGCTC  
49861 GTCATCCCGC CGGACGAGT GCGGTTGAC CCGCCGGGAC TCGCCCGGTG GATGGACGAA  
50 49921 CAGGCGATTA CCGGATCTA CGCGCCGACG GCCGTACTGC GCGCGTGTAT CGAGACGTC  
49981 GATCCGCACA GCGACAGCT CGCCGCCCTG CGGCACCTGT GCCAGGGCGG CGAGGCGCTG  
50041 ATCTCGACG CGCGGTTGCG CGAGCTGTGC CGGCACCGGC CCCACCTGCG CGTGACAAT  
50101 CACTACGGTC CGGCCGAAAG CCAGCTCATC ACCGGGTACA CGCTGCCCGC CGACCCCSAC  
50161 GCGTGGCCCG CCACCGCACC GATCGGCCCG CCGATCGACA ACACCCGCAT CCATCTGCTC  
50221 GACGAGGCGA TCGGGCCGTT TCCGACGCTT ATGCGGGGGC AGCTCTGCGT CGCCGGCGTC  
55 50281 GGCCTCGCCC GTGGGTACCT GGCCCGTCCC GAGCTGACCG CCGAGCGCTG GGTGCCGGGA  
50341 GATGCGGTGCG GCGAGGACG CATGTACCTC ACCGGCGACC TGGCCCGCG CGCGCCGAC  
50401 GGCGACCTG AATTCCTCG CCGGATCGAC GACCAAGTCA AGATCCGCGG CATCCGCGTC  
50461 GACCCGCGTC AGATCGAGAG CCTGCTCGCC GAGGACGCC GCGTACGCA GCGGCGGGT  
50521 TCCGTGCGCG AGGACCGGCG GGGCGAGAAG TTCCTGGCCG CGTACGTCTG ACCGGTGGCC  
60 50581 GGCCGCGACG GCGACGACTT CGCCGCGTGC CTGCGCGCGG GACTGGCCGC CCGGCTGCCC

SUBSTITUTE SHEET (RULE 26)

50641 GCGCGGCTCG TGCCCTCCGC CGTCTCTCTG GTGGAGCGAC TGCCGAGGAC CACGAGCGGC  
50701 AAGGTGGACC GCGCGCGGCT GCGCGAGGCG GAGCGGGGCG CCGGCTCGAC CGGGGCGGTT  
50761 ACGCCCCGCA CCGATGCCGA GCGGAGGCTG TGCCGGATCT TCCAGGAGGT GCTCGACGTC  
50821 GCGCGGCTCG GTGCGGACGA CGACTTCTTC AGGCTCGGCG GCGACTCCCT GCTCGGACCG  
50881 GGGTCTCTCT CCGGCATCCG CGCCGAGGCT GGTGCGGATG TCCGCTCGCG TACGCTCTTC  
50941 GACGGGCGGA CGCCCGCCGC GCTCGCCCTT GCGGCGGACG AGGCCGGGCC GGCCGCCCTG  
51001 CCGCCGATCG CGCCCTCCGC GGAGAACGGG CCGGCCCCCG TCACCGCGGC ACAGGAACAG  
51061 ATGCTGCACT CGCACGGCTC GCTGCTCGCG GCGCCCTCCT ACACGGTCTG CCGGTACGGG  
51121 TTCCGGCTGC GCGGGCCACT CGACCGCGAA GCGCTCGACG CGGACTGAC CCGGATCGCC  
51181 GCGCGCCACG AGCCGCTGCG GACCGGGTTC CCGGATCGGG AACAGGTCGT CCGGCGCGCC  
51241 GCTCCGGTGC GCGCGGAGGT GGTTCGCTTG CCGGTGCGCG ACGTCGACGC CGCGGTCCGG  
51301 GTCGCCCCACC GGGAGCTGAC CCGGCCCTTC GACCTCGTGA ACGGGTCGTT GCTGCGTGCC  
51361 GTGCTGCTGC CGCTGGGCGC CGAGGATCAC GTGCTGCTGC TGATGCTGCA CCACCTCGCC  
51421 GTCTACGGAT GGTCTTCGA CCTCTCTGTC CGGGAGTTGT CCGGGACGCA ACCGGACCTT  
51481 CCGGTGTCTT ACACGGAGT GGCCCTGTTG GAACGGAGTC CGGCCGTGAT CGCGGCCAGG  
51541 GAGAACGACC GGGCCTACTG GCGCCGCGCG CTGGGGGGCG CCACCGCGCC GAGGTGCCC  
51601 GCGGTCCGGC CCGGCGGGGC ACCGACCGGG CCGGCGTTCC TGTGGACGCT CAAGACACC  
51661 GCGGTCTTGG CCGCACGCGG GGTGCGGAC GCGCACGACG CGACGTGCA CGAAACCGTG  
51721 CTGCGCGCCT TCGCCCTGGT CGTGGCGGAG ACCGCGGACA CCGACGACGT GCTCGTCGCG  
51781 ACGCCGTTTC CCGACCGGGG GTACGCGGCG ACCGACCACC TCATCGGCTT CTTGCGGAAG  
51841 GTCCTCGCGC TGCGCTCGA CCTCGGCGCG ACGCGTCTGT TCCCGAGGT GCTGCGCCGG  
51901 GTGCACACCG CGATGGTGGG CGCGGAGGCG CACGAGGCG TGCCCTACTC CGCGCTGCGC  
51961 GCGGAGGACC CCGCGCTGCC GCGGCGCGCG GTGCTGTTCC AGCTCATCAG CGCGCTCAGC  
52021 GCGGAACGTC GGTGCGCGG CATGCAACCG GAGCGGTTCC CCGTCTGTCG CGAGACCGTC  
52081 GACGAGATGA CCGGCGAAGT GTGATGACG CTCTTCGACG ACGGTGCGAC CGTCTCCGGC  
52141 GCGGTGGTCC ACGATGCCGC GCTGCTCGAC CGTGCCACCG TCGACGATTT GCTCACCCTG  
52201 GTGGAGGCGA CGCTGCGTGC CGCGCGGCGC GACCTCACCG TACGCGTCAC CCGTTACGTG  
52261 GAAAGCGAGT AGCCATGCCG GAGCAGGACA AGACAGTCGA GTACCTTCGCG TGGGCGACCG  
52321 CGGAACCTCA GAAGACCCGT GCGGAACCTG CCGGCGACAG CGAGCCGTTG CGGATCGTGG  
52381 GGATGGCCTG CCGGCTGCCC GCGGGGCTCG CGTCGCGGGA GGACCTGTGG CAGTTGCTGG  
52441 AGTCCGGTGG CGACGGCATC ACCGCGTTCC CCACGGACCG GGGCTGGGAG ACCACCGCCG  
52501 ACGGTGCGCG CCGCTTCCTC ACCGGGGCGG CCGGCTTCGA CCGGCGGTTT TTCGGCATCA  
52561 GCGCGCGCGA GCGCTGGCG ATGGACCGCG AGCAGCGCCT GGCCCTGGAG ACCTCGTGGG  
52621 AGGCGTTTCA GCACGCGGGC ATCGATCGCG AGACGCTGCG GGGCAGTGAC ACGGGGGTGT  
52681 TCCTCGGCGC GTTCTTCCAG GGTACGCGCA TCGGCGCGCA CTTGACGGT TACGGGACCA  
52741 CGAGCATTCA CACGAGCGTG CTCTCGGCGC GCCTCGCGTA CTTCTACGGT CTGGAGGGTC  
52801 CCGCGGTGAC GTGCGACAG GCGTGTCTGT CGTCTGCTGT GCGGCTGAC CAGGCGGGGC  
52861 AGTCGCTGCG CTCCGGCGAA TGCTCGCTCG CCCTGGTCTG CCGGCTCACG GTGATGGCCT  
52921 CCGCGCGCGG GTTCGCGGAC TTCTCGGAGC AGGGCGGCGT GGCCCCCGAC GCGCGCTGCA  
52981 AGGCTTCGCG GGAAGCGGCT GACGGCAACG GTTTCGCGGA GGGGTCCGGC GTCTGATCG  
53041 TCGAGAAGCT CTCCGACGCC GAGCGCAACG GCCACGCGT GCTGGCGGTG GTCCGGGGTT  
53101 CCGCGGTCAA CCAGGACGGT GCTTCCGACG GCGGTGTCGC GCGGACCGG CCGTGCAGG  
53161 AGCGGGTGAT CCGGCGAGGC CTGGCCGACG CCGGACTCAC CCGGCGGAC GTGGACCGCG  
53221 TCGAGGCCCA CCGCACCGGC ACCAGGCTGG GCGACCCCAT CGAGGACAG GCCGTGCTGG  
53281 CCACCTACGG GCAGGGGCGC GACACCGCTG TGCTGCTGGG CTCGCTGAAG TCCAACATCG  
53341 GCCACACCCA GCGCGCCGCG GCGCTCGGCG GTGTCATCAA GATGGTCTTC GCCATGCGGC  
53401 ACGGACCCCT GCGCCGACG CTGCACTGCG ACACGCGCTC CTCGACGTC GACTGGACGG  
53461 CCGGCGCGGT CGAACTCTTC ACCGACGCGC GGCCCTGGCC CGAAACCGAC CGCCACGGC  
53521 GCGCGGTTGT CTCCTCTTTC GCGCTGACCG GCACCAACGC CCACATCATC CTCGAAAGCC  
53581 ACCCCCGACC GCGCCCCGAA CCGGCGGCG CACCCGACAC CCGACCGCTG CCGCTGCTGC  
53641 TCTCGGCCCC CACCCCGCAG GCACTCGACG CACAGGTACA CCGCTGCGC GCGTCTCTCG  
53701 ACGACAACCC CCGCGCGGAC CCGGTGCGCG TCGCGCAGAC ACTCGCCCGG CGCACCAGT  
53761 TCGAGCACCG CCGCGTGCTG CTCGGCGACA CGCTCATCAC CGTGAGCCCG AACGCGGGC  
53821 GCGGACCGGT GGTCTTCGTC TACTCGGGGC AAAGCACGCT GCACCCGAC ACCGGGCGGC  
53881 AACTCGCGTC CACTACCCC GTGTCTCGCG AAGCGTGGCG CGAGGCCCTC GACCACCTCG  
53941 ACCCCACCCA GCGCCGCGC AGCACTTTC CCGACAGAC CCGGCTCAC CCGCTCTGCG  
54001 GGTCTGGGG CATCACCCG CACGCGCTCA TCGGCGACTC CCTCGGTGAG ATCAGCGCGC  
54061 GCGACGCGCG GGTGCTCTTC TCCCTGAGCG AGCGGGGCGC GCTCTCACG ACCCGCACCG  
54121 GCTGATGGA CCAACTGCC TCGGGCGCG CGATGGTCA CGTCTGACC AGCGAGGAAA  
54181 AGGCACGCCA GGTGCTGCGC CCGGGCGTGG AGATCGCCCG CCGCAACGGC CCGACTCCC

SUBSTITUTE SHEET (RULE 26)

54241 TCGTGCTGTC CGGGGACGAG GAAGCCGTAC TCGAAGCCGC CCGGCAGCTC GGCATCCACC  
54301 ACCGCCCTGCC GACCCGCCAC GCCGGCCACT CCGAGCCGAT GCAGCCACTC GTCCGCCGCC  
54361 TCGTCGACGT CGCCCGGACC CTGACGTACC ACCAGCCCCA CACCGCCATC CCGGGCGACC  
54421 TCGACACCCG CGAATACTGG GCGCACCAGG TCGCGACCA ACTACGTTTC CAGGGCGACA  
54481 CCGAGCAGTA CCGGGGCGCG ACGTTCCCTCG AGATCGGCCG CAACCAGGAC CTCTCGCCGC  
54541 TCGTCGACGG CGTTCCCGCC CAGACCAGTA CCGCCGACGA GGTGCGGGCG CTGCACACCG  
54601 CGCTCGCGCA GCTCCACGTC CGCGCGCTCG CGATCGACTG GACGCTCGTC CTCGCGGGGG  
54661 ACCGCGCGCC CGTCACGCTG CCCACGTATC CGTTCCAGCA CAAGGACTAC TGGCTGCGGC  
54721 CCACCTCCCG GGCCGATGTG ACCGGCGCGG GGCAGGAGCA GGTGGCGCAC CGCTGCTCG  
54781 GCGCCGCGST CGCGCTGCCC GGCACGGGCG GAGTCGTCCT GACCGGCCCG CTGTGCTGG  
54841 CCTCCCATCC GTGGCTCGGC GAGCACGCGG TCGACGGCAC CGTGCTCCTG CCGGGCGCGG  
54901 CCTTCCTCGA ACTCGCGGCG CGCGCCGGCG ACGAGGTCGG CTGCGACCTG CTGCACGAAC  
54961 TCGTCATCGA GACGCCGCTC GTGCTGCCCC CGACCGGCGG TGTGGCGGTC TCGTCGAGA  
55021 TCGCCGAACC CGACGACAG CGCGGGCGGG CGGTCACCGT CCACGCGCGG GCGGACGGCT  
55081 CGGGCCTGTG GACCCGACAC GCCGGCGGAT TCCTCGGAC CAAGCCGCGG CCGGCCACGG  
55141 CCACGGACCC GGCACCCCTGG CCGCCCGCGG AAGCCGGACC GGTGACGTC CCGGACGCT  
55201 ACGACCGGTT CGAGGACATC GGGTACTCCT ACGGACCGGG CTTCGGGGG CTGCGGGCCG  
55261 CCTGGCGCGC CGCGCACACC GTGTACGCCG AGGTGCGGCT CCGGACGAG CAGAGCGCCG  
55321 ACGCCGCCCG TTTCACGCTG CACCCCGCGC TGCTCGACGC CGCGTTCCAG GCGGGCGCGC  
55381 TCGCCGCGCT CGACGCACCC GCGGGGGCGG CCGGACTGCC GTTCTCGTTT CAGGACGTCC  
55441 GCATCCACGC CGCCGGGCGG ACGCGGTGCG GGTTCACGGT CGGCGCGCAC GCGGAGCGCA  
55501 GCACCGTCCG CATGACCGGC CCGGACGGGC AGCTGGTGGC CGTGGTGGT 3CGGTGCTST  
55561 CCGCCCCCTA CGCGGAAGGC TCCGGTGACG GCCTGCTGCG CCGGTCTGG ACCGAGCTGC  
55621 CGATGCCCTT CCGTCCGCG GACGATCCGC GCGTGGAGGT CCTCGGCGCC GACCCGGGCG  
55681 ACGGCGACGT TCCGGCGGCG ACCCGGGAGC TGACCGCCCG CGTCTCGGC GCGCTCCAGC  
55741 GCCACCTGTC CGCCGCCGAG GACACCACCT TGGTGGTACG GACCGGCACC GGCCCGGCCG  
55801 CTCCCGCCG CGCGGGTCTG GTCCGCTCGG CGCAGGCGGA GAACCCCGCG CGCGTCGTGC  
55861 TCGTCGAGGC GTCCCGGAC ACCTCGGTGG AGCTGCTCGC CGCGTGCGCC GCGCTGGACG  
55921 AACCAGAGCT GGCGTCCGG GACGGCGTGC TCTTCGCGCC CGGCTGGTC CGGATGTCCG  
55981 ACCCCGCGCA CGGCCGCTG TCCCTGCCGG ACGGCGACTG GCTGCTCACC CGGTCCGCT  
56041 CCGGCACGTT GCACGACGTC GCGCTCATAG CCGACGACAC GCGCCGCGG GCGCTCGAAG  
56101 CCGGCGAGGT CCGCATCGAC GTCCGCGCGG CCGGACTGAA CTTCGCGGAT GTGCTGATCG  
56161 CGCTCGGGAC GTACACCGGG GCCACGGCCA TGGGCGGCGA GGCCGCGGGC GTCGTGGTGG  
56221 AGACCGGGCC CGGCGTGGAC GACCTGTCCC CCGGCGACCG GGTGTTCCGG CTGACCCGGG  
56281 CGGCGATCGG CCGGACGGCC GTCACCGACC GCGCGTGGCT GGCCCGGATC CCGGACGGCT  
56341 GGAGCTTCAC CACGGCGGCG TCCGTCCCGA TCGTGTTCGC GACCGCTGG TACGGCTGG  
56401 TCGACCTCGG CACACTGCGC GCGGCGGAGA AGGTCTCTCG CCACGCGGCC ACCGGCGGTC  
56461 TCGGCATGGC CGCCGCACAG ATCGCCCGCC ACCTGGGCGC CGAGCTCTAC GCCACCGCCA  
56521 GTACCGGCA GCAGCACGTC CTGCGCGCCG CCGGGCTGCC CGACACGCAC ATCGCGACT  
56581 CTCGGACGAC CGCGTTCGG ACCGCTTTCC CGCGCATGGA CGTCGTCTTG AACCGGCTGA  
56641 CCGCGAGTT CATCGACGCG TCGCTCGACC TGCTGGACGC CGACGGCCGG TTCGTGAGA  
56701 TGGGCGGCAC CGAGCTGCGC GACCCGACCG CGATCGTCCC CGCTACCTG CGCTCGACC  
56761 TGCTGGACGC GGGCGCCGAC CGCATCGGCG AGATCCTGG CGAACTGCTC CGGTGTTCCG  
56821 ACGCGGGCGC GCTGGAGCCG CTGCCGTTCC GTGCTGSGA CGTCCGGCAG GCACGCGACG  
56881 CGCTCGGCTG GATGAGCCGC GCGCGCCACA TCGGCAAGAA CGTCTGACG CTGCCCGGC  
56941 CGCTCGACCC GGAGGCGGCC GTGCTCTCA CCGGCGGCTC CGGACGCTC GCGGCATCC  
57001 TCGCCCGCCA CCTGCGCGAA CGGCTGTCT ACCTGCTGTC CCGGACGGCA CCGCCCGAGG  
57061 GGACGCCCGG CGTCCACCTG CCCTGCGACG TCGGTGACCG GGACAGCTG GCGCGGGCCC  
57121 TGGAGCGGGT GGACCGGCCG ATCACCAGCG TGGTGCACCT CGCCGGTGC CTGGACGACG  
57181 GCACCGTCGC GTCGCTCACC CCGGAGCGTT TCGACACGGT GCTGCGCCCG AAGGCCGACG  
57241 GCGCCTGGTA CCTGCACGAG CTGACGAAG AGCAGGACCT CGCCGCGTTC GTGCTCTACT  
57301 CGTCGGCCGC CGGCGTGCTC GGCAACGCCG GCCAGGGCAA CTACGTGCGC GCGAACGCGT  
57361 TCGTCGACGC GCTCGCCGAG CTGCGCCACG GTTCCGGGCT GCGGGCCCTC TCCATCGCT  
57421 GGGGCTCTG GGAGGACGTG AGCGGGCTCA CCGCGGCGCT CCGCGAAGCC GACCGGGACC  
57481 GGATCGGCG CAGCGGTTTC CGGGGACATCA CCGCGCAACA GGGCATGCAC CTGTACGAGC  
57541 CGGCCGGCCG CACCGGAAGT CCGGTGGTGG TCGCGGCGGC GCTCGACGAC GCGCGGACG  
57601 TGCGCTGCT GCGCGGCCG CGGCGGACGA CCGTCCGGCG GGCCGCGGTC CCGGAGTGT  
57661 CCGCGGCCG CCGGCTCGCC GCGCTGACCG GCGACGAGCT CCGCGAAGCG CTGCTGACCG  
57721 TCGTCCGGGA GAGCACCGCC GCGGTGCTCG GCCACGTGG TGGCGAGGAC ATCCCGCGCA  
57781 CCGCGGCGTT CAAGGACCTC GGCATCGACT CGTCCACCGC GGTCCAGCTG CCGAACCCCG

57841 TCACCGAGGC GACCGGTGTG CGGCTGAACG CCACGGCGGT CTTCGACTTC CCGACCCCGC  
 57901 ACGTGCTCGC CGGSAAGCTC GGCACGAAC TGACCGGCAC CCGCGCGCCC GTCGTGCCGC  
 57961 GGACCGCGGC CACGGCCGGT GCGCACGACG AGCGGCTGGC GATCGTGGGA ATGGCCTGCC  
 58021 GGCTGCCCGC CGGGSTCGCG TCACCCGAGG AGCTGTGGCA CCTCGTGGCA TCCGGCACCC  
 58081 ACGCCATCAC GGAGTTCCCG ACGGACCGCG GCTGGGACST CGACGCGATC TACGACCCGG  
 58141 ACCCCGACGC GATCGGCAAG ACCTTCGTCC GGCACGGTGG CTTCCTCACC GCGCGACACG  
 58201 GCTTCGACGC GGCCTTCTTC GGCATCAGCC CCGCGGAGGC CCTCGCGATG GACCCGACGC  
 58261 AGCGGGTGCT CCTGGAGACG TCGTGGGAGG CGTTCGAAAG CGCCGGCATC ACCCCGGACT  
 58321 CGACCCGCGC CAGCGACACC GCGGTGTTCC TCGGCGCCTT CTCTACGGT TACGGCACCG  
 58381 GTGCGGACAC CGACGGCTTC GCGCGACCGG GCTCGCAGAC CAGTGTGCTC TCCGGCCGGC  
 58441 TGTCGTACTT CTACGGTCTG GAGGGTCCGG CGGTACGGT CGACACGGCG TGTTCGTCTG  
 58501 CGCTGGTGGC GCTGCACCAG GCCGGGCACT CGCTGCGCTC CGGCGAATGC TCGCTCGCCC  
 58561 TGGTCGGCGG CGTCACGGTG ATGGCGTCTC CCGGCGGCTT CGTGGAGTTC TCCCGGCAGC  
 58621 GCGGCCTCGC GCCGGACGGC CCGGCGAAGG CGTTCGGCGC GGGTGCGGAC GGCACGAGCT  
 58681 TCGCCGAGGG TGCCGGTGTG CTGATCGTCG AGAGGCTCTC CGACGCCGAA CGCAACGGTC  
 58741 ACACCGTCTT GGCGGTCTCG CTGTGTTCCG CGGTCAACCA GGATGGTGCC TCCAACGGGC  
 58801 TGTCGGCGCC GAACGGGCGG GCGCGACCGG GGTGATCCG GCAGGCCCTG GCCAACGGCG  
 58861 GGCTCACCCC GCGGGACGTG GACGCCGTCG AGGCCACCG CACCGGCACC AGGCTGGGCG  
 58921 ACCCCATCGA GGCACAGGCG GTACTGGCCA CCTACGGACA GGAGCGCGCC ACCCCCCTGC  
 58981 TGCTGGGCTC GCTGAAGTCC AACATCGGCC ACGCCAGGC CGCGTCCGGC GTCGCCGGCA  
 59041 TCATCAAGAT GGTGCAGGCC CTCCGGCACG GGGAGCTGCC GCCGACGCTG CACGCCGACG  
 59101 AGCCGTCCGC GCACGTCGAC TGGACGGCGG GCGCCGTGCA ACTGCTGACG TCGGCCCGGC  
 59161 CCGGGCCCGA GACCGACCGG CCACGGCGTG CCGCGGTCTC CTCGTTCGGG GTGAGCGGCA  
 59221 CCAACGCCCA CGTCATCCTG GAGGCGGAC CGGTAACGGA GACGCCCGCG GCATCGCCTT  
 59281 CCGGTGACCT TCCCTGCTG GTGTGCGCAC GCTCACCGGA AGCGCTCGAC GAGCAGATCC  
 59341 GCCGACTGCG CGCCTACCTG GACACCACCC CGGACGTGCA CCGGGTGGCC GTGGCACAGA  
 59401 CGCTGGCCCC GCGCACACAC TTCGCCACCC GCGCCGTGCT GCTCGGTGAC ACCGTCATCA  
 59461 CCACACCCCC CGCGGACCGG CCCGACGAAC TCGTCTTCGT CTACTCCGGC CAGGGCACCC  
 59521 AGCATCCCGC GATGGGCGAG CAGCTCGCCG CCGCCCATCC CGTGTTCGCG GACGCTGGC  
 59581 ATGAAGCGCT CCGCGCCTT GACAACCCCG ACCCCACAGA CCCCACGAC AGCCAGCATG  
 59641 TGCTCTTCGC CCACCAGGCG GCGTTACCG CCCTCCTGCG GTCCTGGGGC ATCACCCTGC  
 59701 ACGCGGTCAT CGGCCACTCG CTGGGCGAGA TCACCGCGGC GCACGCGGCC GGCATCTGT  
 59761 CGCTGGACGA CGCGTGCACC CTGATCACC ACGCGCGCCG CCTCATGCAC ACGTCCCGC  
 59821 CACCCGGTGC CATGGTCACC GTACTGACCA GCGAAGAGAA GGCACGCCAG GCGTTGCGGC  
 59881 CGGGCGTGGA GATCGCCGCC GTCAACGGGG CCCACTCCAT CGTGTGTGCC GGGGACGAGG  
 59941 ACGCCGTGCT CACCGTCGCC GGGCAGTCTG GCATCCACCA CCGCTGCCC GCGCCGACG  
 60001 CCGGGCACTC CGCGCACATG CGCGCGTG GCGCCGAGCT GCTCGCCACC ACCCGCGGGC  
 60061 TCCGCTACCA CCTCCCCAC ACCTCCATTC CGAACGACCC CACCACCGCT GAGTACTGGG  
 60121 CCGAGCAGGT CCGCAAGCCC GTGCTGTTCC ACGCCACGCG GCAGCAGTAC CCGGACGCCG  
 60181 TCTTCGTGGA GATCGGCCCC GCGCAGGACC TCTCCCCCT CGTCGACGGG ATCCCGCTGC  
 60241 AGAACGGCAC CGCGGACGAG GTGCACGCGC TGCACACCGC GCTCGCGCAC CTCTACGCGC  
 60301 GCGGTGCCAC GCTCGACTGG CCGCGCATCC TCGGGGCTGG GTACAGGCAC GACGCGGATG  
 60361 TGCCCGCGTA CGCGTTCCAA CGCGGCCTT ACTGGATCGA GTCGGCACG CCGGCCGAT  
 60421 CCGACGCGGG CCACCCCGTG CTGGGCTCCG GTATCGCCCT CGCCGGGTG CCGGGCGGG  
 60481 TGTTACAGGG TTCTGTGCCG ACCGGTGGCG ACCGCGCGGT GTTCGTGCGC GAGCTGGCGC  
 60541 TGGCCCGCGC GGACGCGGTC GACTGCGCCA CCGTCGAGCG GCTCGACATC CCTCCGTGC  
 60601 CCGGCCGCGC GGGCCATGCC CCGACGACCG TACAGACCTG GGTGACGAG CCGGCGGACG  
 60661 ACGGCCGGCG CCGGTTACCC GTGCACACCC GCACCGGCGA CGCCCGCTGG ACGTGCACG  
 60721 CCGAGGGGGT GCTGCGCCCC CATGGCACGG CCTGCCCGA TGCGGCGCAC GCGGAGTGGC  
 60781 CCCCACCGGG CGCGGTGCC CCGGACGGGC TGCCGGGTGT GTGGCGCGG GGGGACGAGG  
 60841 TCTTCGCGGA GGCCGAGGTG GACGGACCGG ACGGTTTCGT GGTGCACCCC GACCTGCTCG  
 60901 ACGCGGTCTT CTCCGCGGTC GCGGACGGAA GCGCCAGCC GGCCGGATGG CGCGACCTGA  
 60961 CCGTGACGCG GTCGGACGCC ACCGTACTGC GCGCTGCCT CACCCGCGCG ACCGACGGAG  
 61021 CCATGGGATT CGCCGCTTC GACGGCGCCG GCCTGCCGGT ACTACCGCG GAGGCGGTGA  
 61081 CGCTGCGGGA GGTGGCGTCA CCGTCCGGCT CCGAGGAGTC GGACGGCCTG CACCGGTTGG  
 61141 AGTGGCTCGC GGTGCGCGAG GCGGTCTACG ACGGTGACCT GCCCGAGGGA CATGTCCTGA  
 61201 TCACCGCGCG CCACTCCGAC GACCCCGAGG ACATACCCAC CCGCGCGCA ACCCGCGCA  
 61261 CCGCGCTCCT GACCGCCCTG CAACACCACT TCACCACTAC CGACCACTAC CTCATCGTCC  
 61321 ACACCACTAC CGACCCCGCC GCGCGCACCG TCACCGGCT CACCCGCACT GCGGAGAACG  
 61381 AACACCCCA CCGCATCCCG CTCATCGAAA CCGACACCC CCACACCCCT CTCCCCCTGG

SUBSTITUTE SHEET (RULE 26)

	61441	CCCAACTCGC	CACCCTCGAC	CACCCCCACC	TCCGCCTCAC	CCACCACACC	CTCCACCACC
	61501	CCCACCTCAC	CCCCCTCCAC	ACCACCACCC	CACCCACCAC	CACCCCCCTC	AACCCCGAAC
	61561	ACGCCATCAT	CATCACCAGC	GGCTCCGGCA	CCCTCGCCGG	CATCCTCGCC	CGCCACCTGA
5	61621	ACCAACCCCA	CACCTACCTC	CTCTCCCGCA	CCCCACCCCC	CGACGCCACC	CCCGGCACCC
	61681	ACCTCCCTCG	CGACGTCGGC	GACCCCCACC	AACCTCGCCAC	CACCCTCACC	CACATCCCCC
	61741	AACCCCTCAC	CGCCATCTTC	CACACCGCCG	CCACCCTCGA	CGACGGCATC	CTCCACGCCC
	61801	TCACCCCCGA	CGCCCTCACC	ACCGTCTCTC	ACCCCAAAGC	CAACGCCGCC	TGGCACCTGC
	61861	ACCACCTCAC	CCAAAACCAA	CCCCTCACCC	ACTTCGTCTT	CTACTCCAGC	GCCGCCGCCG
	61921	TCCTCGGCAG	CCCCGGACAA	GGAAACTACG	CCGCCGCCAA	CGCCTTCTCT	GACGCCCTCG
10	61981	CCACCCACCG	CCACACCCTC	GGCCACCCCG	CCACCTCCAT	CGCCTGGGGC	ATGTGGCACA
	62041	CCACCAGCAC	CCTCACCAGG	CAACTCGACG	ACGCCGACCG	GGACCGCATC	CGCCCGGGCG
	62101	GTTTCTCTCC	GATCACCAGG	GACGAGGGCA	TGCGCCTCTA	CGAGGCGGCC	GTCGGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCCGGCACA	GCCGATGACC	GGCTCCGTAC
15	62221	CGCCCATCCT	GAGCGGCCGT	CGCAGGAGCG	CGCGGCGCGT	CGCCCGTGCC	GGGCAGACGT
	62281	TCGCCCAGCG	GCTCGCCGAG	CTGCCCGACG	CCGACCGCGG	CGCGGCGCTG	ACCACCCTCG
	62341	TCTCGGACGC	CACGGCCGCC	GTGCTCGGCC	ACGCCGACGC	CTCCGAGATC	CGCCCGACCA
	62401	CGACGTTCAA	GGACCTCGGC	ATCGACTCGC	TCACCGCGAT	CGAGCTGCGC	AACCGGCTCG
	62461	CGGAGGCGAC	CGGGCTGCGG	CTGAGTGCCA	CGCTGGTGTT	CGACCACCCG	ACACCTCGGG
20	62521	TCCTCGCCGC	CAAGCTCCGC	ACCGATCTGT	TCGGCACGGC	CGTGCCACG	CCCGCGCGGA
	62581	CGGCACGGAC	CCACCACGAC	GAGCCACTCG	CGATCGTCGG	CATGGCGTGC	CGACTGCCCG
	62641	GCGGGGTCGC	CTCGCCGGAG	GACCTGTGGC	AGCTCGTGGC	GTCCGGCACC	GACGCGATCA
	62701	CCGAGTTCCC	CACCGACCGC	GGCTGGGACA	TCGACCGGCT	GTTCGACCCG	GACCCGGACG
	62761	CCCCCGGCAA	GACCTACGTC	CGGCACGGCG	GCTTCTCTCG	CGAGGCCGCC	GGCTTCGATG
25	62821	CCGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT	GGACCCGCG	CAGCGCGTCA
	62881	TCCTCGAAAC	CTCCTGGGAG	GCGTTTCGAG	ACGCGGGCAT	CGTGCCGGAC	ACGCTGCGCG
	62941	GCAGCGACAC	CGGCGTGTTT	ATGGGCGCGT	TCTCCCATGG	GTACGGCGCC	GGCGTCGACC
	63001	TGGGCGGGTT	CGGCGCCACC	GCCACGCAGA	ACAGCGTGCT	CTCCGGCCCG	TTGTCTGACT
	63061	TCTTCGGCAT	GGAGGGCCCG	GCCGTACCCG	TCGACACCCG	CTGCTCGTGC	TCGCTGGTCC
30	63121	CCCTGCACCA	GGCGGCACAG	GCGCTGCGGA	CTGGAGAATG	CTCGCTGGCG	CTCGCCGGCG
	63181	GTGTACCGGT	GATGCCACAC	CCGCTGGGCT	ACGTTCGAGT	CTGCCGCCAG	CGGGGACTCG
	63241	CCCCCGACGG	CCGTTGCCAG	GCCTTCGCGG	AAGGCGCCGA	CGGCACGAGC	TTCTCGGAGG
	63301	GCGCCGGCGT	TCTTGTGCTG	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TCGCGGTCTG	CCGCTCCTCC	GCCGTCAACC	AGGACGGCGC	CTCCAACGGC	ATCTCCGCAC
35	63421	CCAACGGCCC	CTCCCAGCAG	CGCGTCATCC	GCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
	63481	CCGCCGACGT	GGACGTGGTG	GAGGCCACCG	GCACCGGAAC	CCCGCTGGGC	GACCCGATCG
	63541	AGGCACAGGC	CATCATCGCG	ACCTACGGCC	AGGACCGCGA	CACACCGCTC	TACCTCGGTT
	63601	CGGTCAAGTC	GAACATCGGA	CACACCCAGA	CCACCGCCGG	TGTCGCCGGC	GTCATCAAGA
	63661	TGGTCATGGC	GATGCGCCAC	GGCATCGCGC	CGAAGACACT	GCACGTGGAC	GAGCCGTCGT
40	63721	CGCATGTGGA	CTGGACCGAG	GGTGCGGTGG	AACTGCTCAC	CGAGGCGAGG	CCGTGGCCCC
	63781	ACGCGGGACG	CCCGCGCCGC	GCGGGCGTGT	CGTCGCTCGG	TATCAGCGGT	ACGAACGCCC
	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	CGCGTGTGGA	GCCGTCTGTT	GACGGGTTGG
	63901	TGCCGTTGCC	GGTGTGCGCT	CGAGTGTAGG	CGAGTCTGCG	GGGGCAGGTG	GAGCGGCTAG
	63961	AGGGGTATCT	GCGCGGGAGT	GTGGATGTGG	CCGCGGTTCG	GCAGGGGTTG	GTGCTGAGC
45	64021	GTGCTGTCTT	CGGTCACCGT	GCGGTACTGC	TGGGTGATGC	CCGGGTGATG	GGTGTGGCGG
	64081	TGGATCAGCC	GCGTACGGTG	TTCGTCTTTC	CCGGGCAGGG	TGCTCAGTGG	GTGGGCATGG
	64141	GTGTGGAGTT	GATGGACCGT	TCTGCGGTGT	TCGCGGCTCG	TATGGAGGAG	TGTGCGCGGG
	64201	CGTTGTTGCC	CCACACGGGC	TGGGATGTGC	GGGAGATGTT	GGCGCGGCCG	GATGTGGCGG
50	64261	AGCGGGTGGA	GGTGGTCCAG	CCGGCCAGCT	GGGCGTTCGC	GGTCAGCCTG	GCCGCACTGT
	64321	GGCAGGCCCA	CGGGGTCTGA	CCCGACGCGG	TGATCGGACA	CTCCAGGGC	GAGATCGCGG
	64381	CGGCGTGCGT	GGCGGGGGCC	CTCAGCCTTG	AGGACGCCGC	CCGCGTGGTG	GCCTTGCGCA
	64441	GCCAGGTCAT	CGCGGCGCGA	CTGGCCGGGC	GGGGAGCGAT	GGCTTCGGTG	GCATTGCCGG
	64501	CCGGTGAGGT	CGGTCTGGTC	GAGGGCGTGT	GGATCGCGGC	GCGTAACGGC	CCCGCCTCGA
	64561	CAGTCGTGGC	CGGCGAGCCG	TCGGCGGTGG	AGGACGTGGT	GACGCGGTAT	GAGACCGAAG
55	64621	GCGTGCGAGT	GCGTCGTATC	GCCGTCGACT	ACGCCTCCCA	CACGCCCCAC	GTGGAAGCCA
	64681	TCGAGGACGA	ACTCGCTGAG	GTAAGGAAGG	GAGTTGCAGG	GAAGGCCGGC	TCGGTGGCGT
	64741	GGTGGTCGAC	CGTGACAGAC	GCCTGGGTGA	CCGAGCCGGT	GGATGAGAGT	TACTGGTACC
	64801	GGAACCTGCG	TGCCCCCGTC	GCGCTGGACG	CGGCGGTGGC	GGAGCTGGAC	GGGTCCGTGT
	64861	TCGTGGAGTG	CAGCGCCCAT	CCGGTGCTGC	TGCCGCGCAT	GGAACAGGCC	CACACGGTGG
60	64921	CGTCGTTGCG	CACCGGTGAC	GGCGGCTGGG	AGCGATGGCT	GACGGCGTTG	GCGCAGGCGT
	64981	GGACCCTGGG	CGCGGCAGTG	GACTGGGACA	CGGTGGTCTGA	ACCGGTGCCA	GGGCGGCTGC

SUBSTITUTE SHEET (RULE 26)

	65041	TCGATCTGCC	CACCTACGCG	TTCGAGCGCC	GGCGCTACTG	GCTGGAAGCG	GCCGGTGCCA
	65101	CCGACCTGTC	CGCGGCCGGG	CTGACAGGGG	CAGCACATCC	CATGCTGGCC	GCCATCACGG
	65161	CACTACCCGC	CGACGACGGT	GGTGTGTGTC	TCACCGGCCG	GATCTCGTTG	CGCACGCATC
5	65221	TCTGGCTGGC	TGATCACGCG	GTGCGGGGCA	CGGTCTCTGT	GCCGGGCACG	GCCCTTGTGG
	65281	AGCTGGTTCAT	CCGGGCCCGT	GACGAGACCG	GTTCGCGGAT	AGTGGATGAA	CTGGTCATCG
	65341	ATCCCCCCT	CGTGGTGCCG	GCGACCGCAG	CCGTGGATCT	GTCGGTGACC	GTGGAAGGAG
	65401	CTGACGAGGC	CGGACGGCGG	CGAGTGACCG	TCCACGCCCC	CACCGAAGGC	ACCGGCAGCT
	65461	GGACCCGGCA	CGCCAGCGGC	ACCCTGACCG	CCGACACCCC	CGACACCCCC	AACGCTTCCG
10	65521	GTGTTGTCCG	TGCGGAGCCG	TTCTCGCAT	GGCCACCTGC	CACTGCCGCG	GCCGTCGACA
	65581	CCTCCGAGTT	CTACTTGCGC	CTGGACGCGC	TGGGCTACCG	GTTCCGACCC	ATGTTCCGCG
	65641	GAATGCGGGC	TGCCTGGCGT	GATGGTGACA	CCGTGTACGC	CGAGGTCGCG	CTCCCGAGG
	65701	ACCGTGCCGC	CGACGCGGAC	GGTTTCGGCA	TGCACCCGGC	GCTGCTCGAC	GCGGCCTTGC
	65761	AGAGCGGCAG	CCTGCTCATG	CTGGAATCGG	ACGGCGAGCA	GAGCGTGCAA	CTGCCGTTCT
15	65821	CCTGGCACGG	CGTCCGGTTC	CACGCGACGG	GCGCGACCAT	GCTGCGGGTG	GCGGTCGTAC
	65881	CGGGCCCGGA	CGGCCTCCGG	CTGCATGCCG	CGGACAGCGG	GAACCGTCCC	GTGCGACGA
	65941	TCGACGCGCT	CGTGACCCGG	TCCCCGGAAG	CGGACCTCGC	GCCCGCCGAT	CCGATGCTGC
	66001	GGGTGCGGGT	GGCCCCGGTG	CCCGTACCTG	CCGGGGCCGG	TCCGTCCGAC	CCGGACGTGC
	66061	TGACGCTGCG	CGGCGACGAC	GCCGACCCGC	TCGGGGAGAC	CCGGGACCTG	ACCACCCGTG
20	66121	TTCTCGACGC	GCTGCTCCGG	GCCGACCGGC	CGGTGATCTT	CCAGGTGACC	GGTGGCCTCG
	66181	CCGCCAAGGC	GGCCGCGAGC	CTGGTCCGCA	CCGCTCAGAA	CGAGCAGCCC	GGCCGCTTCT
	66241	TCCTCGTCGA	AACGGACCCG	GGAGAGGTCC	TGGACGGCGC	GAAGCGCGAC	GCGATCGCGG
	66301	CAGTCGGCGA	GCCCCATGTG	CGGTGCGCG	ACGGCCTCTT	CGAGGCAGCC	CGGCTGATGC
	66361	GGGCCACGCC	GTCCCTGACG	CTCCCGGACA	CCGGGTCTGT	GCAGCTGCGG	CCGTCCGCCA
25	66421	CCGGTTCCCT	CGACGACCTT	GCCGTCGTCC	CCACCGACGC	CCCGGACCCG	CCGTCGCGG
	66481	CCGGCGAGGT	GCGGATCGCG	GTACGCGCGG	CGGGCCTGAA	CTTCCGGGAT	GTCACGGTGC
	66541	CGCTCGGTGT	GGTCGCGGAT	GCGCGTCCGC	TCGGCAGCGA	GGCCGCGGGT	GTCGTCTTGG
	66601	AGACCGGCC	CGGTGTGCAC	GACCTGGCGC	CCGGCGACCG	GGTCTGGGG	ATGCTCGCGG
	66661	GCSCCTTCGG	ACCGGTTCGG	ATCACCGACC	GGCGGTGCT	CGGCCGGATG	CCGGACGGCT
30	66721	GCAGCTTCCC	GCAGGCGGCG	TCCGTGATGA	CCCGCTTCGC	GACCGGTGG	TACGSCCTGG
	66781	TCGACCTGGC	CGGGCTGCGC	CCCGCGGAGA	AGGTCTTGAT	CCACGCGCGG	GCGACCGGTG
	66841	TCGGCGCGGC	GGCCGTCCAG	ATCGCGCGGC	ATCTGGGCGC	GGAGGTGTAC	GCGACCAACA
	66901	GCGCCGCGAA	GCGCCATCTG	GTGGACCTGG	ACGGAGCGCA	TCTGGCCGAT	TCCCGCAGCA
	66961	CCGCGTTCGC	CGACGCGTTC	CCGCGCGTCG	ATGTCGTGCT	CAACTCGCTC	ACCGGTGAAT
35	67021	TCCTCGACGC	GTCCGTCCGG	CTGCTCGCGG	CGGGTGCCCG	GTTCATCGAG	ATGGGGAAGA
	67081	CGGACATCCG	GCACGCCGTC	CAGCAGCCGT	TGCACCTGAT	GGACGCCGGC	CCCGACCGGA
	67141	TGCAGCGGAT	CATCGTCGAG	CTGCTCGGCC	TGTTGCGCGC	CGACGTGCTG	CACCGCTTGC
	67201	CGGTCCACGC	CTGGGACGTG	CGGCAGGCCG	GGGAGGCGTT	CGGCTGGATG	AGCAAGCGGC
	67261	GTCACACCGG	CAAGCTGGTG	CTGACGGTCC	CGCGGCCGCT	GGATCCCGAG	GGGGCCGTGC
40	67321	TCATACCCGG	CGGCTCCGGC	ACCCTCGCCG	GCATCCTCGC	CCGCCACCTG	GGCCACCCCC
	67381	ACACCTACCT	GCTCTCCCGC	ACCCACCCCG	CCGACACCAC	CCCCGGCACC	CACCTCCCTT
	67441	GCGACGTCGG	CGACCCCCAC	CAACTCGCCA	CCACCCTCGC	CCGCATCCCC	CAACCCCTCA
45	67501	CCGCCGTCTT	CCACACCGCC	GGAACCCCTG	ACGACGCCCT	GCTCGACAAC	CTCACCCCGG
	67561	ACCGCGTCGA	CACCGTCTTC	AAACCCAAGG	CCGACGCCGC	CTGGCACCTG	CACCGCTCA
	67621	CCCGCGACAC	CGACCTCGCC	GCGTTCGTGC	TCTACTCCGC	GGTCGCCGGC	CTCATGGGCA
	67681	GCCCGGGGCA	GGGCAACTAC	GTGCGGGCGA	ACGCGTTTCT	CGACGCGCTC	GCCGAACACC
	67741	GCCGTGCGCA	AGGGCTGCCC	GCGCAGTCCC	TCGCATGGGG	CATGTGGGCG	GACGTCAGCG
	67801	CGCTCACC	GAAACTCACC	GACGCGGACC	GCCAGCGCAT	CCGGCGCAGC	GGATTCCCGC
	67861	CSTTGAGCGC	CGCGGACGGC	ATGCGGCTGT	TCGACGCGGC	GACGCGTACC	CCGGAACCGG
50	67921	TCGTCGTGCG	GACGACCGTC	GACCTCACCC	AGCTCGACGG	CGCCGTCGCG	CCGTTGCTCC
	67981	GCGGTCTGGC	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CGCCCGCAAC	GCCGCGGAAG
	68041	AGCCCTTGGC	CGTGCGTCTT	GCCGGGCGTA	CCGCCGCGCA	GCAGCGGCGC	ATCATGAGG
	68101	AGGTCGTGCT	CCGCCACGCG	GCCGCGGTCC	TCGCGTACGG	GCTGGGCGAC	CGCGTGCGCG
	68161	CGGACCGTCC	GTTCCGCGAG	CTCGGTTTCG	ATTGCGTGAC	CGCGGTGAC	CTGCGCAATC
55	68221	GGCTCGCGGC	CGAGACGGGG	CTGCGGCTGC	CGACGACGCT	GGTGTTCAGC	CACCCGACGG
	68281	CGGAGGCGCT	CACCGCCAC	CTGCTCGACC	TGATCGACGC	TCCCACCGCC	CGGATCGCCG
	68341	GGGAGTCCCT	GCCCGCGGTG	ACGGCCGCTC	CCGTGGCGGC	CGCGCGGAC	CCGACGAGC
	68401	CGATCGCCAT	CGTGCGGATG	GCGTGCGGCT	TGTCGCGTGC	CGCGGAGGAC	CGCGGAGGAC
	68461	TGTGGCGGCT	CGTCGAGTCC	GCGACCGACG	CGATCACCAC	GCCTCCTGAC	GACCGCGGCT
60	68521	GGGACGTCGA	CGCGCTGTAC	GACGCGGACC	CGGACGCGGC	CGGCAAGGCG	TACACCTGCG
	68581	GGGGCGGTTA	CCTGGCCGGG	GCGGCGGAGT	TCGACGCGGC	GTTCTTCGAC	ATCATCTCCG

SUBSTITUTE SHEET (RULE 26)



	68641	GCGAAGCGCT	CGGCATGGAC	CCGCAGCAAC	GCCTGCTGCT	CGAAACGGCG	TGGGAGGGGA
	68701	TCGAGCGCGG	CCGGATCAGT	CCGGCGTCGC	TCCGCGGCGG	GGAGGTCGGC	GTCTATGTGG
	68761	GTGCGGCCGC	GCAGGGCTAC	GGGCTGGGCG	CCGAGGACAC	CGAGGGCCAC	GCGATCACCG
5	68821	GTGGTTCCAC	GAGCCTGCTG	TCCGGACGGC	TGGCGTACGT	GCTCGGGCTG	GAGGGCCCGG
	68881	CGGTACACGT	GGACACGGCG	TGCTCGTCGT	CTCTGGTCGC	GCTGCATCTG	GCGTGCCAGG
	68941	GGCTGCGCCT	GGGCGAGTGC	GAACTCGCTC	TGGCCGGAGG	GGTCTCCGTA	CTGAGTTCCG
	69001	CGGCCGCGTT	CGTGGAGTTC	TCCCGCCAGC	GCGGGCTGCG	GGCCGACGGG	CGCTGCAAGT
	69061	CGTTCGGCGC	GGGCGCGGAC	GGCACGACGT	AGTCCGAGGG	CGTGGGCTGC	CTCGTACTGG
10	69121	AACGGCTCTC	CGACGCCGAG	CGGCTCGGGC	ACACCGTGCT	CGCCGTCGTC	CGCGGCACCG
	69181	CCGTACAGTC	CGACGCCGCC	TCCAACGGCC	TCACCGCGCC	GAACGGGCTC	TGCGAGCAGC
	69241	GGGTACATCC	GAAGGCCTGC	GCCGCGGCCG	GGCTGACCGG	CGCCGACGTG	GACGTCGTGG
	69301	AGGGGCACGG	CACCGGCACC	CGGCTCGGGC	ACCCGCTCGA	GGCGGACCGG	CTGCTCGCGA
	69361	GTACGGGGCA	GGACCGTCCG	GCACCGGTCT	GGCTGGGCTC	GCTGAAGTCG	AACATCGGAC
	69421	ATGCCACGGC	CGCGGCCGCT	GTGCGGGGCG	TCATCAAGAT	GGTGCAGGCG	ATCGGCGCGG
15	69481	GCACGATGCC	GCGGACCGTG	CATGTGGAGG	AGCCCTCGCC	CGCGGTCGAC	TGGAGCACCG
	69541	GACAGGTGTC	CCTGCTCGGC	TCCAACGGCC	CCTGGCCGGA	CGCCGTCGCT	CCGCGCCGGG
	69601	CGGCCGTCTC	CGCGTTCGGG	CTCAGCGGGA	CGAACGCGCA	CGTCATCCTG	GAACAGCACC
	69661	GTCCGCGCGC	CGTGGCGTCC	CAGCCGCCCC	GGCCGCCCGG	TGAGGAGTCC	CAGCCGCTGC
	69721	CGTGGGTGCT	CTCCGCGCGG	ACTCCGGCCG	CGCTGCGGGC	CCAGGCGGGC	CGGCTGCGCG
20	69781	ACCACCTCGC	GGCGGCACCG	GACGCGGATC	CGTTGGACAT	CGGGTACCGG	CTGGCCACCA
	69841	GCCCGGCCCA	GTTCCGCCAC	CGTGCCCGGG	TCGTGCGCCG	CACCCCGGAC	GGATTCCGTG
	69901	CCGCGCTCGA	CGGCCTCGCG	GACGGCGCGG	AGGCGCCCGG	AGTCGTACCC	GGGACCGCTC
	69961	AGGAGCGGGC	CGTCGCCTTC	CTCTTCGACG	GCCAGGGCGC	CCAGCGCGCC	GGAAATGGGG
25	70021	GCGAGCTCCA	CCGCCGGTTC	CCCCTCTTCG	CCGCCGCGTG	GGACGAGGTC	TCCGACGCGT
	70081	TCGGCAAGCA	CCTCAAGCAC	TCCCCACCGG	ACGTCTACCA	CGGCGAACAC	GGCGCTCTCG
	70141	CCCATGACAC	CCTGTACGCC	CAGGCCGGCC	TGTTACAGCT	CGAAGTGGCG	CTGCTGCGGC
	70201	TGCTGGAGCA	CTGGGGGGTG	CGGCCGGACG	TGCTCGTCGG	GCACTCCGTC	GGCGAGGTGA
	70261	CCCGCGCGTA	CGCGGCGGGG	GTGCTACCCG	TGGCGGACGG	GACGGAGTTG	ATCGTGGCCC
30	70321	GGGGGCGGGC	GCTGCGGGCG	CTGCCGCCCC	GGGCGATGCT	CGCCGTCGAC	CGCCGCCCCG
	70381	CGGAGGTCCG	CGCCCGCACG	GATCTGGACA	TCGCCGCGGT	CAACGGCCCC	TCCGCCGTGG
	70441	TGCTCGCCGG	TTCGCCGGAC	GATGTGGCGG	CGTTTGAACG	GGAGTGGTCC	GCGGCCGGGC
	70501	GGCGCACGAA	ACGGCTCGAC	GTGCGGCACG	CGTTCCACTC	CCGGCACGTC	GACGGTGGCG
	70561	TCGACGGCTT	CCGTACGGTG	CTGGAGTCGG	TCGCGTTCGG	CGCGGCGCGG	CTGCCGGTGG
35	70621	TGTCCACGAC	GACGGGCCGG	GACGCCGCGG	ACGACCTCAT	AACGCCCGCG	CACTGGCTGC
	70681	GCCATGCGCG	TCGGCCGGTG	CTGTTCTCGG	ATGCCGTCCG	GGAGCTGGCC	GACCGCGCGG
	70741	TCACCACGTT	CGTGGCCGTC	GGCCCTCCCG	GCTCCCTGGC	GTCGGCCCGG	GCGGAGAGCG
	70801	CCGGGGAGGA	CGCCGGGACC	TACCACGCGG	TGCTGCGCGC	CCGACCGGCT	GAGGAGACCG
	70861	CGGCGCTGAC	CGCCCTCGCC	GAGCTGCACG	CCCACGGCGT	CCCGGTGACG	CTGGCCGCGG
	70921	TACTGGCCGG	TGGCCGGCCA	GTGGACCTTC	CCGTGTACGC	GTTCCAGCAC	CGTTCTACTT
40	70981	GGCTGGCCCC	GGCCGTGGCG	GGGGCGCCCG	CCACCGTGGC	GGACACCGGG	GGTCCGGCGG
	71041	AGTCCGAGCC	GGAGGACCTC	ACCGTCGCGG	AGATCGTCCG	TCGGCGCACC	GCGGCGCTGC
	71101	TCGGGCTCAC	GGACCCCGCC	GACGTCGATG	CGGAAGCGAC	GTTCTTCGCG	CTCGGTTTCG
	71161	ACTCACTGGC	GGTGCAGCGG	CTGCGCAACC	AGCTCGCCTC	GGCAACCGGG	CTGGACCTGC
	71221	CGGCGGCCGT	CCTGTTTCGAC	CACGACACCC	CGGCCGCGCT	CACCGCGTTC	CTCCAGGACC
45	71281	GGATCGAGGC	CGGCCAGGAC	CGGATCGAGG	CCGGCGAGGA	CGACGACCGG	CCCACCGTGC
	71341	TCTCGCTCCT	GGAGGAGATG	GAGTCGCTCG	ACGCCGCGGA	CATCGCGGCG	ACGCCGGCCC
	71401	CGGAGCGTGC	GGCCATCGCC	GATCTGCTCG	ACAAGCTCGC	CCATACCTGG	AAGGACTACC
	71461	GATGAGCACC	GATACGCACG	AGGGAACCGG	GCCCGCCCGG	CGCTGCCCAT	TGCGCATCCA
50	71521	GGACGGTCAC	CGCGCCATCC	TGGAGAGCGG	CACGGTGGGT	TCGTTGACCC	TGTTCCGGCT
	71581	CAAGCACTGG	CTGGTTCGCC	CCGCCGAGGA	CGTCAAGCTG	GTCACCAACG	ATCCGCGGTT
	71641	CAGCTCGGCC	GCGCCCTCCG	AGATGCTGCC	CGACCGGCGG	CCCGGCTGGT	TCTCCGGGAT
	71701	GGACTACCCG	GAGCACAACC	GCTACCGCCA	GAAGATCCCG	GGGGACTTCA	CACTGCGGCC
	71761	GGCGCGCAAG	CGGGAGGACT	TCGTGCGCGA	GGCCGCGGAC	GCCTGCGCTG	ACGACATCGA
55	71821	GGCCGCGGGA	CCCGGCACCG	ACCTCATCCC	CGGGTACGCC	AAGCGGCTGC	CCTCCCTCGT
	71881	CATCAACCGG	CTGTACGGGC	TCACCCCTGA	GGAGGGGGCC	GTGCTGGAGG	CACGGATCGG
	71941	CGACATCACC	GGCTCGGCCG	ATCTGCGGAC	CGTCAAGACG	CTGACCGACG	ACTTCTTCGG
	72001	GCACGCGCTG	CGGCTGGTCC	GCGCGAAGCG	TGACGAGCGG	GGCGAGGACC	TGCTGCACCG
	72061	GCTGGCCTCG	GCCGACGACG	GCGAGATCTC	GCTCAGCGAC	GACGAGGCGA	CGGGCGTGTG
	72121	CGCGACGCTG	CTGTTTCGCC	GCCACGACTC	GGTGCAGCAG	ATGGTCCGGT	ACTGCCTCTA
60	72181	CGCACTGCTC	AGCCACCCCG	AGCAGCAGGC	GGCGCTGCGC	GCGCGCCCGG	AGCTGGTCCA

SUBSTITUTE SHEET (RULE 26)



72241 CAACGCGGTC GAGGAGATGC TCCGTTTCCT GCCCGTCAAC CAGATGGGCG TACCGCGCGT  
 72301 CTGTGTCGAG GACGTCGATG TCGGGGGCGT GCGCATCCGT GCGGGGCGACA ACGTGATCCC  
 72361 GCTCTACTCG ACGGCCAACC GCGACCCCGA GGTGTTCCCG CAGCCCGACA CCTTCGATGT  
 72421 GACGCGCCCG CTGGAGGGCA ACTTCGCGTT CGGCCACGGC ATTCACAAGT GTCCCGGCCA  
 5 72481 GCACATCGCC CGGGTGCTCA TCAAGGTGCG CTGCCTGCGG TTGTTGAGC GTTCCCGGA  
 72541 CGTCCGGCTG GCCGGCGACG TGCCGATGAA CGAGGGGCTC GGGCTGTTCA GCCCGGCCGA  
 72601 GCTGCGGGTC ACCTGGGGGG CGGCATGAGT CACCCGGTGG AGACGTTGCG GTTGCCGAAC  
 72661 GGGAGACGG TCGCGCACAT CAACGCGGGG GAGGCGCAGT TCCTCTACCG GGAGATCTTC  
 10 72721 ACCCAGCGCT GCTACCTGCG CCACGGTGTC GACCTGCGCC CGGGGGACGT GGTGTTGAC  
 72781 GTCGGCGCGA ACATCGGCAT GTTCACGCTT TTCGCGCATC TGGAGTGTC TGGTGTGACC  
 72841 GTGCACGCCT TCGAGCCCGC GCCCGTGCCG TTCGCGGCGC TCGGGGCGAA CGTGACGCGG  
 72901 CACGGCATCC CGGGCCAGGC GGACAGTGC GCGGTCTCCG ACAGCTCCGG CACCCGGAAG  
 72961 ATGACCTTCT ATCCCGACGC CACGCTGATG TCCGGTTTC ACGGGATGC CGCGGCCCGG  
 73021 ACGGAGCTGT TGCGCACGCT CGGCCTCAAC GGCGGCTACA CCGCCGAGGA CGTCGACACC  
 15 73081 ATGCTCGCGC AACTGCCCCG CGTCAGCGAG GAGATCGAAA CCCCTGTGGT CCGGCTCTCC  
 73141 GACGTCATCG CGGAGCGCGG TATCGAGGCC ATCGGCCTGC TGAAGGTCGA CGTGGAGAAG  
 73201 AGCGAACGGC AGGTCTTCGC CGCCCTCGAG GACCCGACT GGGCCCGTAT CCGCCAGGTC  
 73261 GTCGCGGAGG TCCACGACAT CGACGGCGCG CTCGAGGAGG TCGTCACGCT GCTCCGCGGC  
 20 73321 CATGGCTTCA CCGTGGTCGC CGAGCAGGAA CCGCTGTTCC CCGGCACGGG CATCCACCAG  
 73381 GTCGCCGCGC GCGGGGTGGC CGGCTGAGCG CCGTCGGGGC CGCGGCCGTC CGCACC GGCG  
 73441 GCGCGGGTGC GGACGGCGGC TCAGCCGGCG TCGGACAGTT CCTTGGGCG TGTGTCAGCG  
 73501 CCCTTACACC CCAGCTTGCG GAACACGTTG GTGAGGTGCT GTTCCACCGT GCTGGAGGTG  
 73561 ACGAACAGCT GGCTGGCGAT CTCCTTGTG GTGCGCCCGA CCGCGGCGTG CGAGGCCACC  
 25 73621 CGCCGCTCCG CCTCGGTGAG CGATGTGATC CGCTGCGCCG GCGTCACGTC CTGGGTCGCG  
 73681 TCCGCGTCCG AGGACTCCCC ACCGAGCCGC CGGAGGAGCG GCACGGCTCC GCACTGGGTC  
 73741 GCGAGGTGCC GTGCGCGGCG GAACAGTCCC CGCGCACGGC TGTGCCGCCG GAGCATGCCG  
 73801 CACGCTTCGC CCATGTGCGC GAGGACGCGG GCCAGCTCGT ACTGGTCGCG GCACATGATG  
 73861 AGCAGATCGG CGGCCTCGTC GAGCAGTTCG ATCCGCTTGG CCGGCGGACT GTAGGCCGCC  
 30 73921 TGCACCCGCA GCGTCATAC CCGCGCCCGG GACCCCATCG GCCGGGACAG CTGCTCGGAG  
 73981 ATGACCTCA GCCCCGTC ACGGCCGCGG CCGAGCAGCA GAAGCGCTTC GCGGCGTCG  
 74041 ACCCGCCACA GGGCCAGGCC CGGCACGTCG ACGGACCAGC GTCGCATCCG CTCGCCGAG  
 74101 TCCCGGAACG CGTTGTACGC CGCCCGGTAC CGCCCGGCCG CGAGATGGTG TTGCCACCG  
 74161 GCCCAGACCA TGTGCAGTCC GAAGAGGCTG TCGGAGGTCT CCTCCGCAA CGGCTCGGCG  
 35 74221 AGCCACCGCT CCGCCCGGTC CAGGTGCGCC AGTCGGATCG CCGCGGCCAC GGTGCTGCTC  
 74281 AGCGGCAATG CGGCGGCCAT CCCCCAGGAG GGCACGACCC GGGGGGCGAG CGCGGCCTCG  
 74341 CCGATTGCA CGGCGGCGGT CAGGTGCGCC CGGCGCAGCG CGGCCTCGG GCGGAACCCC  
 74401 GCGTGGACCG CCTCGTCGCG CGGGGTCGCG ATGTTGTGCT CACCGGCCAG CTGTCGAC  
 74461 CAGGACTGGA CGGCATCGGT GTCCTCGGCG TAGAGCAGGG CCAGCAACGC CATCATGGTC  
 40 74521 GTGGTCCGGT CCGTCGTGAC CCGGGAGTGC TGGAGCACGT ACTCGGCTTT GGCTCGGCC  
 74581 GTTTCGGACC AGCCGCGCAG CGCGTTGCTC AGGGCCTTGT CCGCGACGGC GCGGTGCCGG  
 74641 ACGGCTCCGG AAAACGAGGC GACCTCGTCC TCGGCCGGCG GATCGGCCGG ACGCGGCCGA  
 74701 TCGGCCGCGC CGGGATAGAT CAGCGCGAGG GACAGGTCCG CGACGCGCAG GTGCGCCCGG  
 74761 CCCTGCTCGC TCGGGGCGGC GGAGCGCTGG GCCCGCAGGA CCTCGGCGG CTCGCCCCG  
 45 74821 CGCCCGTCCA TCGCCAGCCA GCAGGCGAGC GACACGGCGT GCTCGCTGGA GAGGACCGT  
 74881 TCCCGCGACG CGGTGAGCAG CTCGGGCACA TGCCGGCCGG ATCTGGCGGG ATCGCAGAGC  
 74941 CGCTCGATGG CGGCGGTGTC GACGCGCAGT GCGGCGTGA CCGCGGGGTC GTCGGAGGCC  
 75001 CGGTAGGCGA ACTCCAGGTA GGTGACGGCC TCGTCGAGCT CGCCGCGCAG GTGGTGCTCG  
 75061 CGCGCGGCGT CGGTGAACAG CCCGGCGACC TCGGCGCGCT GCACCCGGCC GGTACCCATC  
 50 75121 TGGTGGCGGG CGAGCACCTT GCTGGCCACG CCGCGGTCCC GCAGCAGTTA CAGCGCCAGC  
 75181 TCGTCAGGC CACGCCGCTC GCGGCGGGAG AGGTGCTGCA GTACGACGGA GCGGGCCGCG  
 75241 GGGTGGCGGA ACCGCCCTTC CCGCAGCAGC CGCCCTCGA CCAGCTGTTT GTGGCCCTGC  
 75301 TCGACCGCCT CGGTGTCGAG GCCGGTCATC CGCTGGACGA GGGTGAGTTC GACACTCTCG  
 75361 CCGAGCACGG CGGAAGCTCG GGCGACGCTC AGCGCGGCGG GGCCGCAACG ATAGAGCGAC  
 75421 CCGAGGTAGG CGAGCCGGTA CGCCCGCCCC GCGACCACTT CCAGGCACCC TGAGGTCCGT  
 55 75481 GTCCGTGCCT CCCGGATGTC GTCGATCAGG CCGTGGCCGA GGAGCAGGTT GCCCGCGGTC  
 75541 CCGCGGAACG CCTGGGCCAC CAGTCTGTCG TCGCGTCTCT GGCCGAGGTG CCGGCGCACG  
 75601 AGTTCGTTGG TCTGCGCTC GGTGAGCGGG CGCAGCGCGA TCTCTGTGTA GTGCGCAGA  
 75661 CTCAGCAGTG CCGCCCGGAA TTGGGAGTGG GCGGGCGTCG GCCGAGGCA CTCGGTCAGC  
 75721 ACGATGGCGA CACGGGCCCG GCTGATGCGG CGCGCGAGGT GGAGCAGGCA GCGCAGCGAC  
 60 75781 GCGCGTCCG CGTGGTGCAC GTCGTCGATG CCGATCAGTA CCGGCGGCTC CGCGGCGAGC

SUBSTITUTE SHEET (RULE 26)

75841 GTCAGCACCG TCGGGGTGAG TTCGGTCCCC AGGCGGTTGT CGACGTCGGC CGGCAGGTTT  
 75901 TCGCACGATG CCGTCAGCCG GACCAGCTCC GGTGTCCGGG CGGCCAGCTC GGGCTGGTCG  
 75961 AGGAGCTGGC CGAGCATGCC GTACGGCAGG GCCCCTCCT CCATGGAGCA CACCGCGCGA  
 5 76021 AGGGTGACGA AGCCGGCCTT GGCCGCGCGC GCGTCGAGGA GTTCGGTCTT GCCCGAGGCG  
 76081 ATCGGCCCCG TGACGGCGGC GACGACGCCC CGCCCGCCCC CCGCTCGGGT GAGCGCCCCG  
 76141 TGGAGGGAAC CGAACTCGTC ATCGCGGGCG ATCAGGTCTG GGGGAGATAA GCGCGCTATC  
 76201 ACGAATGGAA CTACCTCGCG ACCGTCTGGG AAACCCATAG GCATCACATG GCTTGTGTGAT  
 76261 CTGTACGGCT GTGATTGAGC CTGGCGGGAT GCTGTGCTAC AGATGGGAAG ATGTGATCTA  
 10 76321 GGGCCGTGCC GTTCCCTCAG GAGCCGACCG CCCCCGGCGC CACCCGCCGT ACCCCTGGG  
 76381 CCACCAGCTC GGCGACCCGC TCCTGGTGGT CGACGAGGTA GAAGTGCCCG CCGGGGAAGA  
 76441 CCTCCACCGT GGTCGGCGCG GTCGTGTGCC CGGCCCAGGC GTGGGCCTGC TCCACCGTCG  
 76501 TCTTCGGATC GTCGTACCG ATGCACACCG TGATCGGCGT CTCCAGCGGC GGCGCGGGCT  
 76561 CCCACCGGTA CGTCTCCGC CGTAGTAGT CCGCCCGCAA CGGCGCCAGG ATCAGCGCGC  
 76621 GCATTTCTGTC GTCCGCCATC ACATCGGCGC TCGTCCCGCC GAGGCCGATG ACCGCCGCCA  
 15 76681 GCAGCTCGTC GTCGGACGCG AGGTGGTCCT GGTCGGCGCG CGGCTGCGAC GCGGCCCGCC  
 76741 GGCCCGAGAC GATCAGGTGC GCCACCGGGA GCCGTGGGC CAGCTCGAAC GCGAGTGTCT  
 76801 CGCCCATGCT GTGGCCGAAC AGCACCAGCG GACGTTCCAG CCCCCGCTTC AACGCCTCGG  
 76861 CCACGAGGCC GGCGAGAACA CGCAGGTCGC GCACCGCCTC CTCGTGCGCG CGSTCCTGGC  
 76921 GGCCGGGGTA CTGCACGGCG TACACGTCCG CCACCGGGGC GAGCGCACCG GCCAGCGGAA  
 20 76981 GGTAGAACGT CGCCGATCCG CCGCGTGGG GCAGCAGCAC CACCCGTACC GGGGCCTCGG  
 77041 GCGTGGGGAA GAACTGCCGC AGCCAGAGTT CCGAGCTCAC CGCACCCCTC CCGCCGCGAC  
 77101 CTGGGGAGCC CGGAACCGGG TGATCTCGGC CAAGTGCTTC TCCCGCATCT CCGGTCTGGT  
 77161 CACGCCCCAT CCCTCCTCCG GCGCCAGACA GAGGACGCCG ACTTTGCCGT TGTGCACATT  
 77221 GCGATGCACA TCGCGCACCG CCGACCCGAC GTCGTGAGC GGGTAGGTCA CCGACAGCGT  
 25 77281 CGGGTGCACC ATCCCCTTGC AGATCAGGCG GTTCGCCTCC CACGCCTCAC GATAGTTCGC  
 77341 GAAGTGGGTA CCGATGATCC GCTTCACGGA CATCCACAGG TACCGATTGT CAAAGGCGTG  
 77401 CTCGTATCCC GAGGTTGACG CGCAGGTGAC GATCGTGCCA CCCCACGTG TCACGTAGAC  
 77461 ACTCGCGCCG AACGTCGCGC GCCCGGGTG CTGAACACG ATGTCGGGAT CGTCACCGCC  
 30 77521 GGTACAGTCC CGGATC

Those of skill in the art will recognize that, due to the degenerate nature of the  
 genetic code, a variety of DNA compounds differing in their nucleotide sequences can be  
 used to encode a given amino acid sequence of the invention. The native DNA sequence  
 encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to  
 35 illustrate a preferred embodiment of the invention, and the present invention includes DNA  
 compounds of any sequence that encode the amino acid sequences of the polypeptides and  
 proteins of the invention. In similar fashion, a polypeptide can typically tolerate one or more  
 amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or  
 significant loss of a desired activity. The present invention includes such polypeptides with  
 40 alternate amino acid sequences, and the amino acid sequences shown merely illustrate  
 preferred embodiments of the invention.

The recombinant nucleic acids, proteins, and peptides of the invention are many and  
 diverse. To facilitate an understanding of the invention and the diverse compounds and  
 methods provided thereby, the following general description of the FK-520 PKS genes and  
 45 modules of the PKS proteins encoded thereby is provided. This general description is  
 followed by a more detailed description of the various domains and modules of the FK-520

PKS contained in and encoded by the compounds of the invention. In this description, reference to a heterologous PKS refers to any PKS other than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

The FK-520 PKS is composed of three proteins encoded by three genes designated *fkbA*, *fk bB*, and *fk bC*. The *fk bA* ORF encodes extender modules 7 - 10 of the PKS. The *fk bB* ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The *fk bC* ORF encodes extender modules 5 - 6 of the PKS. The *fk bP* ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520 polyketide.

The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihydroxycyclohexene carboxylic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin, FK-506, rifamycin, and avermectin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction

with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide.

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA, a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is merely added to coding sequences for modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous

PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

5 In an illustrative embodiment of this aspect of the invention, the invention provides recombinant PKSs and recombinant DNA compounds and vectors that encode such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS encoded thereby can be fed or supplied with N-acylcysteamine thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

15 The second extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

25 In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these

replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence  
5 can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for  
10 malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS.  
15 The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA  
20 compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the  
25 malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding  
30 sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an

FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

5 The fourth extender module of the FK-520 PKS includes a KS, an AT that binds ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS.  
10 The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fourth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a  
15 DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA  
20 specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of  
25 the FK-520 PKS, a PKS for a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender  
30 module of the FK-520 PKS.

As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS

genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzymes of the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl-21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA



specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-functional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-19 to C-20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds

of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the sixth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as,

for example, the coding sequences for extender module two encoded by the *eryAI* gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or

malonyl CoA specific AT; deleting the KR, the DH, and/or the ER; and/or replacing the KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-15-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

In another illustrative embodiment, the present invention provides a hybrid PKS in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the eighth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such

analog is preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth  
5 extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13-desmethoxy) FK-506  
10 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds  
15 of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of  
20 the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the  
25 FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific  
30 AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can

originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, and an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the tenth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a

module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

5 The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the *fkbP* gene and so provides recombinant methods for expressing the *fkbP* gene product in recombinant host cells. The recombinant *fkbP* genes of the invention include those in which the coding  
10 sequence for the adenylation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of pipecolic acid (see  
15 Nielsen *et al.*, 1991, *Biochem.* 30: 5789-96). The *fkbL* gene encodes a homolog of RapL, a lysine cyclodeaminase responsible in part for producing the pipecolate unit added to the end of the polyketide chain. The *fkbB* and *fkbL* recombinant genes of the invention can be used in heterologous hosts to produce compounds such as FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel polyketides and non-ribosomal peptides.

20 The present invention also provides recombinant DNA compounds that encode the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520. Figure 2 shows the various sites on the FK-520 polyketide core structure at which these enzymes act. By providing these genes in recombinant form, the present invention provides recombinant host cells that can produce FK-520. This is accomplished by introducing the recombinant  
25 PKS, P450 oxidase, and methyltransferase genes into a heterologous host cell. In a preferred embodiment, the heterologous host cell is *Streptomyces coelicolor* CH999 or *Streptomyces lividans* K4-114, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference. In addition, by providing recombinant host cells  
30 that express only a subset of these genes, the present invention provides methods for making FK-520 precursor compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds and vectors that are useful in generating, by homologous recombination, recombinant host



cells that produce FK-520 precursor compounds. In this aspect of the invention, a native host cell that produces FK-520 is transformed with a vector (such as an SCP2\* derived vector for *Streptomyces* host cells) that encodes one or more disrupted genes (i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those genes.

5 When the vector integrates by homologous recombination, the native, functional gene is deleted or replaced by the non-functional recombinant gene, and the resulting host cell thus produces an FK-520 precursor. Such host cells can also be complemented by introduction of a modified form of the deleted or mutated non-functional gene to produce a novel compound.

10 In one important embodiment, the present invention provides a hybrid PKS and the corresponding recombinant DNA compounds that encode those hybrid PKS enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that comprises all or part of one or more modules and thioesterase/cyclase domain of a first PKS and all or part of one or more modules, loading module, and thioesterase/cyclase domain of a second PKS.

15 In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapamycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonyl CoA.

25 In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specific for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60/091,526, incorporated herein by reference.

5 The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

(i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS,

15 but also:

(ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,

20 (iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520 PKS genes, and

(iv) from combinations of the foregoing.

Various hybrid PKSs of the invention illustrating these various alternatives are described herein.

25 Examples of the production of a hybrid PKS by co-expression of PKS genes from the FK-520 PKS and another non-FK-520 PKS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples include (i)

30 replacement of the *fkfC* gene with the *rapB* gene; and (ii) replacement of the *fkfA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell

is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily modified to act only as neurotrophins, as described in Example 6, below.

5 Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the *fkba* gene of an FK-520 or FK-506 producing host cell with a hybrid *fkba* gene in which:  
(a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the rapamycin PKS; and  
(b) the module 8 coding sequences have been replaced by the module 8 coding sequence of  
10 the rifamycin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13-desmethoxy-13-methyl-FK-520 or 13-desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the  
15 producing host cell by a vector such as pHU204, which is a plasmid pRM5 derivative that has the well-characterized SCP2\* replicon, the *colE1* replicon, the *tsr* and *bla* resistance genes, and a *cos* site. This vector can be used to introduce the recombinant *fkba* replacement gene in an FK-520 or FK-506 producing host cell (or a host cell derived therefrom in which the endogenous *fkba* gene has either been rendered inactive by  
20 mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a  
25 module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau *et al.*, 1999, "Dissecting the role of acyltransferase domains of  
30 modular polyketide synthases in the choice and stereochemical fate of extender units," *Biochemistry* 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau *et al.*, *supra*. One can

also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale *et al.*, 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," *Science* 284: 482-485, incorporated herein by reference.

5       The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the present invention.

10   **Avermectin**

U.S. Pat. No. 5,252,474 to Merck.

MacNeil *et al.*, 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

15       MacNeil *et al.*, 1992, *Gene* 115: 119-125, Complex Organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase.

Ikedo *et al.*, Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA* 96: 9509-9514.

20   **Candicidin (FR008)**

Hu *et al.*, 1994, *Mol. Microbiol.* 14: 163-172.

**Epothilone**

U.S. Pat. App. Serial No. 60/130,560, filed 22 April 1999.

**Erythromycin**

25       PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5,824,513 to Abbott.

Donadio *et al.*, 1991, *Science* 252:675-9.

Cortes *et al.*, 8 Nov. 1990, *Nature* 348:176-8, An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of

30   *Saccharopolyspora erythraea*.

Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

**FK-506**

Motamedi *et al.*, 1998, The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506, *Eur. J. biochem.* 256: 528-534.

Motamedi *et al.*, 1997, Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur. J. Biochem.* 244: 74-80.

#### Methyltransferase

US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from *Streptomyces* MA6858. 31-O-desmethyl-FK-506 methyltransferase.

Motamedi *et al.*, 1996, Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and FK-520, *J. Bacteriol.* 178: 5243-5248.

#### ***Streptomyces hygroscopicus***

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

#### **Lovastatin**

U.S. Pat. No. 5,744,350 to Merck.

#### **Narbomycin**

U.S. patent application Serial No. 60/107,093, filed 5 Nov. 1998, and Serial No. 60/120,254, filed 16 Feb. 1999.

#### **Nemadectin**

MacNeil *et al.*, 1993, *supra*.

#### **Niddamycin**

Kakavas *et al.*, 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

#### **Oleandomycin**

Swan *et al.*, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano *et al.*, 1998, Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

#### **Picromycin**

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue *et al.*, 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pikC*-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chemistry & Biology* 5(11): 661-667.

- 5 Xue *et al.*, Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity, *Proc. Natl. Acad. Sci. USA* 95: 12111 12116.

#### **Platenolide**

EP Pat. App. Pub. No. 791,656 to Lilly.

#### **Rapamycin**

- 10 Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA* 92:7839-7843.

Aparicio *et al.*, 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene* 169: 9-16.

15 **Rifamycin**

August *et al.*, 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

#### **Sorangium PKS**

- 20 U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

#### **Soraphen**

U.S. Pat. No. 5,716,849 to Novartis.

- Schupp *et al.*, 1995, *J. Bacteriology* 177: 3673-3679. A *Sorangium cellulosum* (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen  
25 A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

#### **Spiramycin**

U.S. Pat. No. 5,098,837 to Lilly.

#### Activator Gene

- 30 U.S. Pat. No. 5,514,544 to Lilly.

#### **Tylosin**

EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

Kuhstoss *et al.*, 1996, *Gene* 183:231-6.. Production of a novel polyketide through the construction of a hybrid polyketide synthase.

#### Tailoring enzymes

Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five  
5 tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the  
10 FK-520 PKS in PCT patent publication No. 98/51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073,538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the  
15 hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT  
20 domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived for example, from the picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS  
25 enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two  
30 carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08/989,332, filed 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cell/vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.



The present invention provides a wide variety of expression vectors for use in *Streptomyces*. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2\* (see Hopwood *et al.*, *Genetic Manipulation of Streptomyces: A Laboratory manual* (The John Innes Foundation, Norwich, U.K., 1985); Lydiate *et al.*, 1985, *Gene* 35: 223-235; and Kieser and Melton, 1988, *Gene* 65: 83-91, each of which is incorporated herein by reference), SLP1.2 (Thompson *et al.*, 1982, *Gene* 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth *et al.*, 1989, *Mol. Gen. Genet.* 219: 341-348, and Bierman *et al.*, 1992, *Gene* 116: 43-49, each of which is incorporated herein by reference), or a high copy number vector, such as pIJ101 and pJV1 (see Katz *et al.*, 1983, *J. Gen. Microbiol.* 129: 2703-2714; Vara *et al.*, 1989, *J. Bacteriol.* 171: 5782-5781; and Servin-Gonzalez, 1993, *Plasmid* 30: 131-140, each of which is incorporated herein by reference). Generally, however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an *E. coli* origin of replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood *et al.*, *supra*).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin), *tsr* (confers resistance to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter, typically with an attendant ribosome binding site sequence. The present invention provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the *fkfO* gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the *fkfO* and *fkfB* genes. The *fkfO* promoter is believed to be bi-directional in that it promotes transcription of the genes *fkfO*, *fkfP*, and *fkfA* in one direction and *fkfB*, *fkfC*, and *fkfL* in the other. Thus, in one aspect, the present invention

provides a recombinant expression vector comprising the promoter of the *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO*. In a preferred embodiment the transcribed gene is an FK-520 PKS gene. In another preferred embodiment, the transcribed gene is a gene that encodes a protein comprised in a hybrid PKS.

5 Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the *actI* promoter and its attendant activator gene *actII-ORF4*, which is provided in the pRM1 and pRM5 expression vectors, *supra*. This promoter is activated in the stationary phase of growth when secondary metabolites are normally  
10 synthesized. Other useful *Streptomyces* promoters include without limitation those from the *ermE* gene and the *melC1* gene, which act constitutively, and the *tipA* gene and the *merA* gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to *Streptomyces* and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is  
15 inserted into a neutral site of the chromosome or in a vector under the control of the inducible *merA* promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the *actII-ORF4* gene discussed above include *dnrI*, *redD*, and *ptpA* genes (see U.S. patent application Serial No. 09/181,833,  
20 *supra*) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the  
25 biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl CoA.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes are sufficient to confer this ability on *Streptomyces* host cells. For conversion of 2-  
30 hydroxymalonyl to 2-methoxymalonyl, the *fkbG* gene is also employed. While the complete coding sequence for *fkbH* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence

herein shows one T, there may be two, resulting in an extension of the *fkbH* reading frame to encode the amino acid sequence:

MTIVKCLVWDLNLTWRGTVLEDDEVVLTDEIREVITTLDDRGILQAVASKNDHD  
 LAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERA EVA  
 5 FHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRLMYQAGFARDQAREAYSGPD  
 EDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRALLTDP AHE  
 VLVVTMGDRFGPHGAVGIILLEKKPSTWHLKLLATSCRVVVSFGAGATILNWLTDQG  
 ARAG AHLVADFRRTDRNRMM EIAYRFAGFADSDCPCVSEVAGASAAGVERLHLEP  
 SARPAPTTTLTAADIAPVTVSAAG.

10 For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkbS* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkbE* and *fkbU* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the  
 15 recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in  
 20 Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesize ethylmalonyl CoA or 2-hydroxymalonyl CoA.  
 25 The invention provides methods and vectors for constructing recombinant *Streptomyces coelicolor* and *Streptomyces lividans* that are able to synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

In a preferred embodiment, the present invention provides recombinant  
 30 *Streptomyces* host cells, such as *S. coelicolor* and *S. lividans*, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that

comprise one or more AT domains specific for ethylmalonyl CoA. Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in which one or more AT domains is specific for ethylmalonyl CoA.

5 In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For example, deletion or inactivation of the *fkfG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the  
10 corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkfG* gene product acts on 2-hydroxymalonyl and the resulting 2-methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-  
15 506) are replaced by either hydrogen or methyl.

This possibility of non-specific binding results from the construction of a hybrid PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520  
20 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13,15-didesmethoxy-FK-520 and corresponding FK-506 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis, i.e., the *fkfH*, *fkfI*, *fkfJ*, and *fkfK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for  
25 ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the resulting host cells are thus preferred for production of polyketides that do not require the same.

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g., U.S.  
30 Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-

desmethoxy-FK-520; 13,15-didesmethoxy-FK-506; 13,15-didesmethoxy-FK-520; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-520; 13,15-didesmethoxy-18-hydroxy-FK-506; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent  
5 Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,466; and 4,920,218, incorporated herein by reference.

Other compounds of the invention are shown in Figure 8, Parts A and B. In Figure 8, Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration  
10 and are applied to the dermis for treatment of conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in Figure 8, Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group,  
15 where R<sub>3</sub> and R<sub>4</sub> can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure 8, Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15-methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-  
20 32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of  
25 Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or triazole derivatives provides the C-32 tetrazole or triazole derivative. As shown in the lower scheme of Figure 8, Part B, reacting the starting compound with p-nitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

30 The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active

ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, talc, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically, parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg

to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the present invention may be administered on an intermittent basis, i.e., at semi-weekly, weekly, semi-monthly, or  
5 monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate  
10 and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of, for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8%  
15 by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with  
20 compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex, and diet of the subject; the  
25 time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

A detailed description of the invention having been provided above, the following examples are given for the purpose of illustrating the present invention and shall not be  
30 construed as being a limitation on the scope of the invention or claims.

#### Example 1

#### Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-520

The C-13 methoxyl group is introduced into FK-520 via an AT domain in extender module 8 of the PKS that is specific for hydroxymalonyl and by methylation of the hydroxyl group by an S-adenosyl methionine (SAM) dependent methyltransferase.

5 Metabolism of FK-506 and FK-520 primarily involves oxidation at the C-13 position into an inactive derivative that is further degraded by host P450 and other enzymes. The present invention provides compounds related in structure to FK-506 and FK-520 that do not contain the C-13 methoxy group and exhibit greater stability and a longer half-life *in vivo*. These compounds are useful medicaments due to their immunosuppressive and neurotrophic activities, and the invention provides the compounds in purified form and as pharmaceutical compositions.

10 The present invention also provides the novel PKS enzymes that produce these novel compounds as well as the expression vectors and host cells that produce the novel PKS enzymes. The novel PKS enzymes include, among others, those that contain an AT domain specific for either malonyl CoA or methylmalonyl CoA in module 8 of the FK-506 and FK-520 PKS. This example describes the construction of recombinant DNA compounds that encode the novel FK-520 PKS enzymes and the transformation of host cells with those recombinant DNA compounds to produce the novel PKS enzymes and the polyketides produced thereby.

To construct an expression cassette for performing module 8 AT domain replacements in the FK-520 PKS, a 4.6 kb *Sph*I fragment from the FK-520 gene cluster was cloned into plasmid pLitmus 38 (a cloning vector available from New England Biolabs). The 4.6 kb *Sph*I fragment, which encodes the ACP domain of module 7 followed by module 8 through the KR domain, was isolated from an agarose gel after digesting the cosmid pKOS65-C31 with *Sph* I. The clone having the insert oriented so the single *Sac*I site was nearest to the *Spe*I end of the polylinker was identified and designated as plasmid pKOS60-21-67. To generate appropriate cloning sites, two linkers were ligated sequentially as follows. First, a linker was ligated between the *Spe*I and *Sac*I sites to introduce a *Bgl*II site at the 5' end of the cassette, to eliminate interfering polylinker sites, and to reduce the total insert size to 4.5 kb (the limit of the phage KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5'-CTAGTGGGCAGATCTGGCAGCT-3'  
3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.



Next, a linker of the following sequence was ligated between the unique *Sph*I and *Afl*III sites of plasmid pKOS60-27-1 to introduce an *Nsi*I site at the 3' end of the module 8 cassette. The linker employed was:

5' -GGGATGCATGGC-3'  
 3' -GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (*Avr* II or *Nhe* I) and 3' end (*Xho* I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers *Spe*Bgl-fwd and either *Avr*-rev or *Nhe*-rev:

*Spe*Bgl-fwd 5' -CGACTCACTAGTGGGCAGATCTGG-3'

*Avr*-rev 5' -CACGCCTAGGCCGGTCGGTCTCGGGCCAC-3'

*Nhe*-rev 5' -GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 µl reaction, 5 µl of 10x *Pfu* polymerase buffer (Stratagene), 5 µl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 µl DMSO, 2 µl of each primer (10 µM), 1 µl of template DNA (0.1 µg/µl), and 1 µl of cloned *Pfu* polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (*Bgl*II and *Avr*II or *Spe*I and *Nhe*I), and cloned into either pLitmus 28 or pLitmus38 (New England Biolabs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2, respectively.

Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers *Bsr*Xho-fwd and *Nsi*Afl-rev:

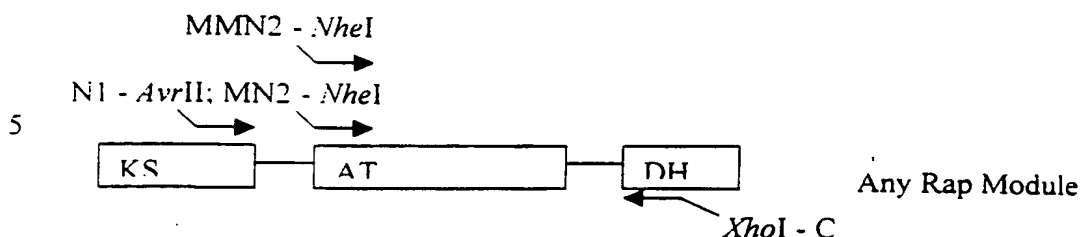
*Bsr*Xho-fwd 5' -GATGTACAGCTCGAGTCGGCACGCCCCGGCCGCATC-3'

*Nsi*Afl-rev 5' -CGACTCACTTAAGCCATGCATCC-3'

PCR conditions were as described above. The PCR fragment was cut with *Bsr*GI and *Afl*III, gel isolated, and ligated into pKOS60-37-4 cut with *Asp*718 and *Afl*III and inserted into pKOS60-37-2 cut with *Bsr*GI and *Afl*III, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *Avr*II and *Xho*I or *Nhe*I and *Xho*I, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *AvrII* or *NheI* site at the 5' end and an *XhoI* site at the 3' end. The PCR conditions were as given above and the primer sequences were as follows:

- 5  
RATN1 5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3'  
(3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA),  
RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3'  
(Rap AT shorter version 5'- sequence and specific for malonyl CoA),  
10 RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3'  
(Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and  
RATC 5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3'  
(Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).



Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned PCR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The *AvrII*-*XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

```

20 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
   I W Q L A E A L L T L V R E S T
   GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
   A A V L G H V G G E D I P A T A A
   GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
   F K D L G I D S L T A V Q L R N
   CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
   A L T E A T G V R L N A T A V F D
   TTCCCGACCCCGCACGTGCTCGCCGGAAGCTCGGCGACGAAGTACCGG 250
   F P T P H V L A G K L G D E L T G
   CACCCGCGCGCCCCGTGTCGCCCGGACCGCGGCCACGGCCGGTGCGCACG 300
   T R A P V V P R T A A T A G A H
   ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350
   D E P L A I V G M A C R L P G G V
   GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
   A S P E E L W H L V A S G T D A I
   CACGGAGTTCCCGACGGACCGCGGTGGGACGTGACGCGATCTACGACC 450
   T E F P T D R G W D V D A I Y D
   CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
   P D P D A I G K T F V R H G G F L
   ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550
   T G A T G F D A A F F G I S P R E
   GGCCCTCGCGATGGACCCGACGAGCGGGTGTCTCTGGAGACGTCTGTGGG 600
   A L A M D P Q Q R V L L E T S W
   AGGCGTTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650
   E A F E S A G I T P D S T R G S D
   ACCGGCGTGTTCGTGCGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700
   T G V F Y G A F S Y G Y G T G A D
   CACCGACGGCTTCGGGCGGACCGGCTCGCAGACCAAGTGTGCTCTCCGGCC 750
   T D G F G A T G S Q T S V L S G
50 GGCTGTCTGTAATTCTACGGTCTGGAGGGTCCGGCGGTACGGTCCGACACG 800

```

R L S Y F Y G L E G P A V T V D T  
 CGGTGTTCTGCTGCTGGTGGCGCTGCACCAGSGCGGGGCTGCTGCG 850  
 A C S S S L V A L H Q A G Q S L R  
 CTGCGCGGATGCTGCTGCGCGCTGCTGCGCGGCTGACGCTGATGGCGT 900  
 S G E C S L A L V G G V T V M A  
 CTGCGCGGCGCTGCTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950  
 S P G G F V E F S R Q R G L A P D  
 GGCGCGGCGAAGGCGTTTCGGCGCGGCTGCGGACGCGAGCTTCGCCGA 1000  
 G R A K A F G A G A D G T S F A E  
 GGGTGCCGCTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
 G A G V L I V E R L S D A E R N  
 GTCACACCGTCTTGGCGGTGCTCCGTGGTTTCGGCGGTCAACCAGGATGGT 1100  
 G H T V L A V V R G S A V N Q D G  
 GCCTCCAACGGGTGTTCGGCGCCGACGGGCGCTCGCAGGAGCGGGTGAT 1150  
 A S N G L S A P N G G P S Q E R V I  
 CCGGCAGGCGCTTGGCCAACGCCGGGCTCACCCCGCGGACGTGGACGCC 1200  
 R Q A L A N A G L T P A D V D A  
 TCGAGGCCCCACGGCACCGGCACCGAGGCTGGGCGACCCCATCGAGGCACAG 1250  
 V E A H G T G T R L G D P I E A Q  
 GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300  
 A V L A T Y G Q E R A T P L L L G  
 CTCGCTGAAGTCCAACATCGGCCACGCCCGCGCTCGCGCTCGCGG 1350  
 S L K S N I G H A Q A A S G V A  
 GCATCATCAAGATGCTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
 G I I K M V Q A L R H G E L P P T  
 CTGCACGCCGACGAGCCGTGCGCGCACGTGACTGGACGGCGCGCGCGCT 1450  
 L H A D E P S P H V D W T A G A V  
 CGAAGTGTGACGTGCGGCCCGCGGTGGCCCGAGACCGACCGGCCTAGGC 1500  
 E L L T S A R P W P E T D R P R  
 GGGCAGGCGTGTGCTCTTCCGGTACGTGGCACCAACGCCACCGTCATC 1550  
 R A G V S S F G I S G T N A H V I  
 CTGGAAAGCGCACCCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG 1600  
 L E S A P P T Q P A D N A V I E R  
 GGCACCGGAGTGGGTGCCGTTGGTGATTTCCGGCCAGGACCCAGTCGGCTT 1650  
 A P E W V P L V I S A R T Q S A  
 TGACTGAGCAGGAGGGCGGTTGCGTGCGTATCTGGCGCGCTCGCCCGGG 1700  
 L T E H E G R L R A Y L A A S P G  
 GTGGATATGCGGGCTGTGGCATCGACGCTGGCGATGACACGGTGGTGTT 1750  
 V D M R A V A S T L A M T R S V F  
 CGAGCACCGTGGCGTGTGCTGGGAGATGACACCGTCACCGGCACCGCTG 1800  
 E H P A V L L G D D T V T G T A  
 TGTCTGACCTCGGGCGGTGTTGCTCTTCCCGGGACAGGGGTGCGAGCGT 1850  
 V S D P R A V F V F P G Q G S Q R  
 GCTGGCATGGGTGAGGAAGTGGCCGCGGCTTCCCGCTCTTCGCGCGGAT 1900  
 A G M G E E L A A A F P V F A R I  
 CCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACG 1950  
 H Q Q V W D L L D V P D L E V N  
 AGACCGGTTACGCCACGCCGGCCCTGTTTCGAATGCAGGTGGCTCTGTTT 2000  
 E T G Y A Q P A L F A M Q V A L F  
 GGGCTGTGGAATCGTGGGTGTACGACCGGACGCGGTGATCGGCCATT 2050  
 G L L E S W G V R P D A V I G H S  
 GGTGGGTGAGCTTGGCGCTGCGTATGTGTCCGGGTGTGGTCTGTTGGAGG 2100  
 V G E L A A A Y V S G V W S L E  
 ATGCCTGCACTTTGGTGTGGCGCGGGCTCGTCTGATGCAGGCTCTGCCC 2150  
 D A C T L V S A R A R L M Q A L P  
 CGGGGTGGGGTGATGGTCTGCTCCCGGTCTCGGAGGATGAGGCCCGGGC 2200  
 A G G V K V A V P V S E D E A R A  
 CGTGTGGGTGAGGTTGGGAGATCGCCCGGTCAACGGCGCGCTGCTGCG 2250  
 V L G E G V E I A A V N G P S S  
 TCGTTCTCTCCGGTGATGAGGCCCGCGGTGCTGCAGGCCGCGGAGGGGCTC 2300

**SUBSTITUTE SHEET (RULE 26)**

V V L S G D E A A V L Q A A E S L  
GGGAAGTGGACGCGGCTGGCGACCAGCCACGCGTTCCATTCCGCCCGSTAT 2350  
G K W T R L A T S H A F H S A F M  
5 GGAACCCATGCTGGAGGASTTCCGGGCGGTCCCGGAAGGCGTGCCTACC 2400  
E P M L E E F P A V A E G L T Y  
GGACGCCGCGAGGTCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAG 2450  
E T P Q V S M A V G D Q V T T A E  
TACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGC 2500  
Y W V R Q V R D T V R F G E Q V A  
10 CTCGTACGAGGACGCCGTGTTTCGTGCGAGCTGGGTGCCGACCGGTCACTGG 2550  
S Y E D A V F V E L G A D R S L  
CCCGCCTGGTGCAGCGGTGTCCGATGCTGCACGGCGACCACGAAATCCAG 2600  
A R L V D G V A M L H G D H E I Q  
GCCGCGATCGGCGCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCGA 2650  
15 A A I G A L A H L Y V N G V T V D  
CTGGCCCCGCGCTCCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTC 2700  
W P A L L G D A P A T R V L D L  
CGACATACGCCCTTCCAGCACCAGCGCTACTGGCTCGAGTCGGCAGCGCCG 2750  
P T Y A F Q H Q R Y W L E S A R P  
20 GCCGCATCCGACGCGGGCCACCCCGTGTGGGCTCCGGTATCGCCCTCGC 2800  
A A S D A G H P V L G S G I A L A  
CGGGTCGCGCGGGCGGGTGTTCACGGGTTCGGTGGCGACCGGTGCGGACC 2850  
G S P G R V F T G S V P T G A D  
GCCGCGGTGTTTCGTGCGCGAGCTGGCGCTGGCCGCGCGGACGCGGTGCGAC 2900  
25 R A V F V A E L A L A A D A V D  
TGCGCCACGGTCGAGCGGCTCGACATCGCCTCCGTGCCCCGCGCGCGGG 2950  
C A T V E R L D I A S V P G R P G  
CCATGGCCGGACGACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGACG 3000  
H G R T T V Q T W V D E P A D D  
30 GCCGGCGCGCGTTTACCGTGCACACCCGCACCGGCGACGCCCCGTGGACG 3050  
G R R F T V H T R T G D A P W T  
CTGCACGCGGAGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCCGATGC 3100  
L H A E G V L R P H G T A L P D A  
GGCCGACGCCGAGTGGCCCCCACCAGGGCGCGGTGCCCCGCGGACGGGCTGC 3150  
35 A D A E W P P P G A V P A D G L  
CGGGTGTGTGGCGCCGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGAC 3200  
P G V W R R G D Q V F A E A E V D  
GGACCGGACGGTTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTC 3250  
G P D G F V V H P D L L D A V F S  
40 CGCGGTGCGCGACGGAAGCCGCCAGCCGGCCGGATGCGCGACCTGACGG 3300  
A V G D G S R Q P A G W R D L T  
TGCACGCGTCGGACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCC 3350  
V H A S D A T V L R A C L T R R T  
GACGGAGCCATGGGATTCCCGCCTTCGACGGCGCGGCTGCGGCTACT 3400  
45 D G A M G F A A F D G A G L P V L  
CACCGCGGAGGCGGTGACGCTGCGGGAGGTGGCGTACCGTCCGGCTCCG 3450  
T A E A V T L R E V A S P S G S  
AGGAGTCGGACGGCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCG 3500  
E E S D G L H R L E W L A V A E A  
50 GTCTACGACGGTGACCTGCCGAGGGACATGTCCTGATCACCGCCGCCCA 3550  
V Y D G D L P E G H V L I T A A H  
CCCCGACGACCCGAGGACATACCCACCGCGCCACACCCGCGCCACCC 3600  
P D D P E D I P T R A H T R A T  
GCGTCTTGACCGCCCTGCAACACCACCTCACCACCACCGACCACACCCCTC 3650  
55 R V L T A L Q H H L T T T D H T L  
ATCGTCCACACCACCGACCCCGCGGCGCCACCGTCACCGGCCTCAC 3700  
I V H T T T D P A G A T V T G L T  
CCGACCCGCGGAGAACGACACCCCGCGGATCCGCGCTCATCGAAACCG 3750  
R T A Q N E H P H R I R L I E T  
60 ACCACCCCGACACCCCGCTCCCGCTGGCCCAACTCGCCACCCCTCGACCAC 3800

SUBSTITUTE SHEET (RULE 26)

D H P H T P L P L A Q L A T L D H  
 CCCCACCTCCGGCTCACCACCCACACCCCTCCACCACCCCCACCTCACCCC 3850  
 P H L R L T H H T L H H P H L T P  
 CCTCCACACCCACCCACCCACCCACCCACCCCTCAACCCCGAACACG 3900  
 5 L H T T T P P T T T P L N P E H  
 CCATCATCATCAGGGGGCTCCGGCACCCCTCGCCGGCATCCTCGCCCGC 3950  
 A I I I T G G S G T L A G I L A R  
 CACCTGAACCAACCCACACCTACCTCCTCTCCCGCACCCACCCCGCA 4000  
 H L N H P H T Y L L S R T P P P D  
 10 CCCCACCCCGGGACCCACCTCCCTGCGACGTCCGGCGACCCCAACCAAC 4050  
 A T P G T H L P C D V G D P H Q  
 TCGCCACCAACCTCACCACATCCCCAACCCCTCACCGCCATCTTCCAC 4100  
 L A T T L T H I P Q P L T A I F H  
 ACCGCCGCCACCCCTCGACGACGGCATCCTCCACGCCCTCACCCTCGACCG 4150  
 15 T A A T L D D G I L H A L T P D R  
 CCTCACCACCGTCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACC 4200  
 L T T V L H P K A N A A W H L H  
 ACCTCACCCAAACCAACCCCTCACCACCTTGGTCTCTACTCCAGCGCC 4250  
 H L T Q N Q P L T H F V L Y S S A  
 20 GCCGCCGTCTCGGACGCCGAGCAAGGAACTACGCCGCCGCAACGC 4300  
 A A V L G S P G Q G N Y A A A N A  
 CTTCTCGACGCCCTCGCCACCCACCGCCACACCTCGGCCAACCCGCCA 4350  
 F L D A L A T H R H T L G Q P A  
 CCTCCATCGCCTGGGGCATGTGGCACACCACCAGCACCTCACCAGGACAA 4400  
 25 T S I A W G M W H T T S T L T G Q  
 CTGACGACGCCGACCGGGACCGCATCCGCCGCGGCGGTTTCTCCCGAT 4450  
 L D D A D R D R I R R G G F L P I  
 CACGGACGACGAGGGCATGGGGATGCAT  
 30 T D D E G

The *AvrII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

35 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 Q L A E A L L T L V R E S T  
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100  
 A A V L G H V G G E D I P A T A A  
 GTTCAAGCACTCGGCATCGACTCGCTCACCAGCGGTCCAGCTGCGCAACG 150  
 40 F K D L G I D S L T A V Q L R N  
 CCCTCACCAGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250  
 F P T F H V L A G K L G D E L T G  
 45 CACCCGCGCGCCCGTCTGTGCCCCGGACCGCGGCCACGGCCGGTGGCGCAG 300  
 T R A P V V P R T A A G A H  
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGCGGGGTC 350  
 D E P L A I V G M A C R L P G G V  
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400  
 50 A S P E E L W H L V A S G T D A I  
 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450  
 T E F P T D R G W D V D A I Y D  
 CGGACCCCGACGCGATCGGCAAGACCTTCTGTCGGGCACGGTGGCTTCCTC 500  
 P C P D A I G K T F V R H G G F L  
 55 ACCGGCGCGACAGGCTTCGACGCGCGTCTCTCGGCATCAGCCGCGCGA 550  
 T G A T G F D A A F F G I S P R E  
 GGCCCTCGCGATGGACCCGAGCAGCGGGTGTCTCTGGAGACGTCTGTTGG 600

A L A M D P Q Q R V L L E T S W  
AGGCGTTTCGAAAGGCGCGGCATCACCCCGGACTCGAACC CGGCGCAGCGAC 650  
E A F E S A G I T P D S T R G S D  
ACCGGGGTGTTCTCGCGCGCTTCTCTACGGTTACCGCACCGGTGCGGA 700  
5 T G V F Y S A F S Y G Y G T G A D  
CACCGACGGCTTCCGCGCGACCGGCTCGCAGACCACTGTGCTCTCGCGCC 750  
T D G F E A T G S Q T S V L S G  
GGCTGTCTACTTCTACGGTCTGGAGGGTCCGCGCGTACGGTCCACACG 800  
R L S Y F Y G L E G P A V T V D T  
10 CCGTGTCTGCTCTGCTGGTGGCGCTGCACCAGGCGCGGCGAGTCTGCTGCG 850  
A C S S S L V A L H Q A G Q S L R  
CTCCGGCGAATGCTCGCTCGCCCTGGTCTGGCGGGCGTCACGGTGATGGCGT 900  
S G E C S L A L V G G V T V M A  
CTCCCGCGGGCTTCTGTTGAGTTCTCCCGGCAGCGCGGCTCTGCGCCGGAC 950  
15 S P G G F V E F S R Q R G L A P D  
GGCCGGGCGAAGGCGTTTCGCGCGGGTTCGGACGGCACGAGCTTCGCGCA 1000  
G R A K A F G A G A D G T S F A E  
GGGTGCCGGTGTCTGATCTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
G A G V L V E R L S D A E R N  
20 GTCACACCGTCTCGCGGTCTGCTCGTGGTTCGCGCGTCAACCAGGATGGT 1100  
G H T V L A V V R G S A V N Q D G  
GCCTCCAACGGGCTGTCTCGCGCGCGAACGGGCGGTCTCGAGAGCGGGTGAT 1150  
A S N G L S A P N G P S Q E R V I  
CCGCGAGGCCCTCGCGAACGCGGGGCTCACCCCGGCGGACGTGGACGCCG 1200  
25 R Q A L A N A G L T P A D V D A  
TCGAGGCCCCACGGCACCGGACCGAGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G T G T R L G D F I E A Q  
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300  
A V L A T Y G Q E R A T P L L L G  
30 CTCGCTGAAGTCCAACATCGGCCACGCCCAGGCCGCGTCCGGCGTCTGCGG 1350  
S L K S N I G H A Q A A S G V A  
GCATCAAGATGGTGCAGGCCCTCCGGCACGGGAGCTGCCGCCGACG 1400  
G I I K M V Q A L R H G E L P P T  
CTGCACGCCGACGAGCCGTCGCGCACGTCGACTGGACGGCGCGCGCGT 1450  
35 L H A D E P S P H V D W T A G A V  
CGAACTGCTGACGTGCGCCCCGGCGTGGCCCCGAGACCGACCGCCTAGGC 1500  
E L L T S A R P W P E T D R P R  
GGGCGGGCGTGTCTCTCTCGGAGTCAGCGGCACCAACGCCCACGTCATC 1550  
R A G V S S F G V S G T N A H V I  
40 CTGGAGAGCGCACCCGCCGCTCAGCCCCGCGGAGGAGCGCAGCCTGTTGA 1600  
L E S A P P A Q P A E E A Q P V E  
GACGCCGGTGGTGGCCTCGGATGTGCTGCGGCTGGTGATATCGGCCAAGA 1650  
T P V V A S D V L P L V I S A K  
CCGAGCCCCCGCTGACCGAACACGAAGACCGGCTGCGCGCCTACCTGGCG 1700  
45 T Q P A L T E H E D R L R A Y L A  
GCGTCCGCCCGGGCGGATATACGGGCTGTGGCATCGACGCTGGCGGTGAC 1750  
A S P G A D I R A V A S T L A V T  
ACGGTCCGGTGTTCGAGCACCGCGCCGTA CTCTTGAGATGACACCGTCA 1800  
R S V F E H R A V L L G D D T V  
50 CCGGCACCGCGGTGACCGACCCAGGATCGTGTTTGTCTTTCCCGGGCAG 1850  
T G T A V T D P R I V F V F P G Q  
GGGTGGCAGTGGCTGGGATGGGAGTGCAGTGCAGGATTCTGTCGGTGGT 1900  
G W G L G M G S A L R D S S V V  
GTTCCGCGAGCGGATGGCCGAGTGTGCGGCGGCGTTGCGCGAGTTCGTGG 1950  
55 F A E R M A E C A A A L R E F V  
ACTGGGATCTGTTACGGTCTGAGTATCCGGCGGTGGTGGACCGGGTT 2000  
D W E L F T V L D D P A V V D R V  
GATGTGGTCCAGCCGCTCTCTGGGCGATGATGGTTTCCCTGGCGCGGT 2050  
D V V Q P A S W A M M V S L A A V  
60 GTGGCAGGCGCGCGGTGTCGGCGGATGCGGTGATCGGCCATTTCGCAGG 2100

W Q A A G V R P D A V I G H S Q  
GTGAGATCGCCGAGCTTGTGTGGCGGGTGCGGTGTCACGCGATGCC 2150  
G E I A A A C V A G A V S L R D A  
5 GCGCGGATCGTGACCTTCCGAGCCAGGCGATCGCCCGGGGCTGCGCGG 2200  
A R E V T L R S Q A I A R G L A G  
CGCGGGCGGATGSCATCCGTGCGCCCTGCCCGCGCAGGATGTGAGCTGG 2250  
R G A M A S V A L P A Q D V E L  
TCGACGGGGGCTGGATCGCCGCCACACGCGCGCCGCTCCACCTGATC 2300  
V D G A W I A A H N G P A S T V I  
10 GCGGGCACCCCGGAAGCGGTGACCATGTCTCACCCTCATGAGSCACA 2350  
A G T P E A V D H V L T A H E A Q  
AGGGGTGCGGGTGCGGGCGGATCACCGTCGACTATGCCTCGCACACCCCGC 2400  
G V R V R R I T V D Y A S H T P  
15 ACCTCGAGCTGATCCGCGACGAACACTCGACATCACTAGCGACAGCAGC 2450  
H V E L I R D E L L D F T S D S S  
TCGCGACACCCGCTCGTGCCGTGGCTGTGACCGTGGACGGCACCTGGGT 2500  
S Q T P L V P W L S T V D G T W V  
CGACAGCCCGCTGGACGGGGAGTACTGGTACCGGAACCTGCGTGAACCGG 2550  
D S P L D G E Y W Y R N L R E P  
20 TCGGTTTCCACCCCGCGTCAGCCAGTTGCAGGCCAGGGCGACACCGTG 2600  
V G F H P A V S Q L Q A Q G D T V  
TTGCTCGAGGTCAGCGCCAGCCCGGTGTTGTTGCAGGCGATGGACGACGA 2650  
F V E V S A S P V L L Q A M E D D  
TGTCGTCACGGTTCCACGCTGCGTGTGACGACGGCGACGCCACCCGGA 2700  
25 V V T V A T L R R D D G D A T R  
TGCTCACCGCCCTGGCACAGGCCTATGTCCACGGCGTCACCGTCGACTGG 2750  
M L T A L A Q A Y V H G V T V D W  
CCCGCCATCCTCGGACACCACCAACCCGGGTACTGGACCTTCCGACCTA 2800  
P A I L G T T T T T R V L D L F T Y  
30 CGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGGCACGCCCGGCCGCAT 2850  
A F Q H Q R Y W L E S A R P A A  
CCGACGCGGGCCACCCCGTGTCTGGGCTCCGGTATCGCCCTCGCCGGGTG 2900  
S D A G H P V L G S G I A L A G S  
CCGGGCGGGGTGTTACGGGTTCGGTGCCGACCGGTGCGGACCGCGCGGT 2950  
35 P G R V F T G S V P T G A D R A V  
GTTCTGTCGCGAGCTGGCGCTGGCCGCGCGGACGCGGTGACTGCGCCA 3000  
F V A E L A L A A A D A V D C A  
CGGTGAGCGGCTCGACATCGCCTCCGTGCCCCGCGCGCGGGCCATGGC 3050  
T V E R L D I A S V P G R P G H G  
40 CGGACGACCGTACAGACCTGGGTGACGAGCCGGCGGACGACGGCCGGCG 3100  
R T T V Q T W V D E P A D D G R R  
CCGTTTACCGTGCACACCCGACCGGCGACGCCCGTGGACGCTGCACG 3150  
R F T V H T R T G D A P W T L H  
CCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGCGGCCGAC 3200  
45 A E G V L R P H G T A L P D A A D  
GCCGAGTGGCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGCCGGGTGT 3250  
A E W P P P G A V P A D G L P G V  
GTGGCGCCGGGGGACAGGTCTTCGCCGAGGCGGAGGTGGACGGACCGG 3300  
W R R G D Q V F A E A E V D G P  
50 ACGGTTTCTGTTGACACCCGACCTGCTCGACGCGGTCTTCTCCGCGGTC 3350  
D G F V V H P D L L D A V F S A V  
GGCGACGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGC 3400  
G D G S R Q P A G W R D L T V H A  
GTCGGACGCCACCGTACTGCGCGCTGCCTCACCGGGCGCACCGACGGAG 3450  
55 S D A T V L R A C L T R R T D G  
CCATGGGATTGCCCGCTTCGACGGCGCCGCTGCCGGTACTCACCGCG 3500  
A M G F A A F D G A G L P V L T A  
GAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTC 3550  
E A V T L R E V A S P S G S E E S  
60 GGACGCCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACG 3600



D G L H R L E W L A V A E A V Y  
 ACGGTGACCTGCCCGAGGGACATGTCTGATCACCAGCCGCCACCCCGAC 3650  
 D G D L P E G H V L I T A A H P D  
 GACCCCGAGGACATACCCACCCCGCCGACACCCCGCCGCGTCT 3700  
 5 C P E D I P T R A H T R A T R V L  
 GACCCGCTGCAACACCACCTCACCACCACCGACCACACCCCTCATCGTCC 3750  
 T A L Q H H L T T T D H T L I V  
 ACACCACCACCGACCCCGCCGCGCCACCGTCACCGGCCTCACCAGCACC 3800  
 H T T T D P A G A T V T G L T R T  
 10 CCCCAGAACGAACACCCACCGCATCCGCTCATCGAAACCGACCACCC 3850  
 A Q N E H P H R I R L I E T D H P  
 CCACACCCCTCCCTGGCCCACTCGCCACCCTCGACCACCCCCACC 3900  
 H T P L P L A Q L A T L D H P H  
 TCCGCTCACCACACACCTCCACCACCCACCTCACCCTCCAC 3950  
 15 L R L T H H T L H H P H L T P L H  
 ACCACCACCCACCCACCCACCCCTCAACCCCGAACACGCCATCAT 4000  
 T T T P P T T T P L N P E H A I I  
 CATACCGGCGCTCCGCGACCTCGCCGGCATCTCGCCGCCACCTGA 4050  
 I T G G S G T L A G I L A R H L  
 20 ACCACCCACACCTACCTCTCTCCGACCCACCCCGACGCCACC 4100  
 N H P H T Y L L S R T P P P D A T  
 CCGGCGACCCACCTCCCTGCGACGTGCGGACCCCACTCGCCAC 4150  
 P G T H L P C D V G D P H Q L A T  
 CACCTCACCACATCCCAACCCCTCACCAGCATCTTCCACACCGCG 4200  
 25 T L T H I P Q P L T A I F H T A  
 CCACCCTCGACGACGGCATCTCCACGCCCTCACCCTCGACCGCTCACC 4250  
 A T L D D G I L H A L T P D R L T  
 ACGTCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACCACCTCAC 4300  
 T V L H P K A N A A W H L H H L T  
 30 CCAAAACCAACCCCTCACCCTCTCGTCTCTACTCCAGCGCCGCCGCG 4350  
 Q N Q P L T H F V L Y S S A A A  
 TCCTCGGCGACCCCGGACAAGGAACTACGCCGCGCAACGCCTTCCTC 4400  
 V L G S P G Q G N Y A A A N A F L  
 GACGCCCTCGCCACCCACCGCCACACCTCGGCCAACCCGCCACCTCCAT 4450  
 35 D A L A T H R H T L G Q P A T S I  
 CGCCTGGGGCATGTGGCACACCAGCACCTCACCAGCAACTCGACG 4500  
 A W G M W H T T S T L T G Q L D  
 ACGCCGCGGGACCGCATCCGCGCGGCGTTTCTCCCGATCACGGAC 4550  
 D A D R D R I R R G G F L P I T D  
 40 GACGAGGGCATGGGGATGCAT  
 D E G

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with  
 the endogenous AT domain replaced by the AT domain of module 12 (specific for malonyl  
 45 CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence  
 shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 Q L A E A L L T L V R E S T  
 GCGCCGCTGCTCGGCCACGTGGGTGGCGAGGACATCCCGCGACGGCGGC 100  
 50 A A V L G H V G G E D I P A T A A  
 GTTCAAGGACCTCGGCATCGACTCGCTCACCAGCGGTCCAGCTGCGCAACG 150  
 F K D L G I D S L T A V Q L R N  
 CCCTCACCAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 55 TTCCCGACCCCGACGTGCTCGCCGGGAAGCTCGGCGACGAAGTACCGG 250  
 F P T P H V L A G K L G D E L T G  
 CACCCGCGCGCCGCTCGTCCCGGACCGCGGCCACGGCCGGTGGCAGC 300

T R A F V V P R T A A T A G A H  
ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350  
D E P L A I V G M A C R L P G G V  
GGGTCAACCCGAGGAGCTGTGGCACTCGTGGCATCCGGCACCGACGCCAT 400  
5 A S P E E L W H L V A S G T D A I  
CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450  
T E F F T D R G W D V D A I Y D  
CGGACCCCGACGGGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500  
F D P D A I G K T F V R H G G F L  
10 ACCGGCGCGACAGGCTTCGACGCGGGCTTCTTCGGCATCAGCCCGCGCGA 550  
T G A T G F D A A F F G I S P R E  
GGCCCTCGCGATGGACCCGCGAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600  
A L A M D P Q R V L L E T S W  
AGGCGTTCGAAAGCGCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650  
15 E A F E S A G I T P D S T R G S D  
ACCGGCGTGTTTCGTCCGGCGCTTCTCCTACGGTTACGGCACCGGTGCGGA 700  
T G V F V G A F S Y G Y G T G A D  
CACCGACGGCTTCGGCGCGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750  
T D G F G A T G S Q T S V L S G  
20 GGCTGTCGTAATTCTACGGTCTGGAGGGTCCGGCGGTACGGTCGACACG 800  
R L S Y F Y G L E G P A V T V D T  
CGGTGTTTCGTCTCGTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850  
A C S S S L V A L H Q A G Q S L R  
CTCCGGCGGAATGCTCGCTCGCCCTGGTCCGGCGGCTCACGGTGATGGCGT 900  
25 S G E C S L A L V G G V T V M A  
CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTTCGCGCCGGAC 950  
S P G F V E F S R Q R G L A P D  
GGCCGGGCGAAGGCGTTCCGGCGGGTGCAGGACGGCACGAGCTTCGCCGA 1000  
G R A K A F G A G A D G T S F A E  
30 GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
G A G V L I V E R L S D A E R N  
GTCACACCGTCCTGGCGGTCTCGTGGTTCGGCGGTCAACCAGGATGGT 1100  
G H T V L A V V R G S A V N Q D G  
GCCTCCAACGGGCTGTCCGGCGCCGAACGGGCGTTCGAGGAGCGGGTGAT 1150  
35 A S N G L S A P N G P S Q E R V I  
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGGACGTGGACGCCG 1200  
R Q A L A N A G L T P A D V D A  
TCGAGGCCCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250  
V E A H G T G T R L G D P I E A Q  
40 GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300  
A V L A T Y G E R A T P L L L G  
CTCGCTGAAGTCCAACATCGGGCACGCCCAGGCCGCTCCGGCGTCCGCCG 1350  
S L K S N I G H A Q A A S G V A  
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
45 S I I K M V Q A L R H G E L P P T  
CTGCACGCCGACGAGCCGTCGCCGCACGTCGACTGGACGGCCGGCGCCGT 1450  
L H A D E P S P H V D W T A G A V  
CGAFACTGCTGACGTCCGGCCCGGCGTGGCCCGAGACCGACCGGCCACGGC 1500  
E L L T S A R P W P E T D R P R  
50 GTSCCGCGCTCTCCTCGTTCCGGGTGAGCGGCACCAACGCCACGTCATC 1550  
R A A V S S F G V S G T N A H V I  
CTGSAGGCCGGACCGGTAACGGAGACGCCCCGCGGCATCGCCTTCGGGTGA 1600  
L E A G P V T E T P A A S P S G D  
CCTTCCCTGCTGGTGTCCGGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650  
55 L P L V S A R S P E A L D E Q  
TCTGCCGACTGCGCGCCTACCTGGACACACCCCGGACGTCGACCGGGTG 1700  
T R R L R A Y L D T T P D V D R V  
GGGTTGGCAGACGCTGGCCCGGGCGCACACACTTCGCCCCACCGCGCGT 1750  
A V A Q T L A R R T H F A H R A V  
60 GCTGCTCGGTGACACCGTCATCACACACCCCCCGGACCGGGCCCGACG 1800

SUBSTITUTE SHEET (RULE 26)

L L G E T V I T T P P A D R P D  
AACTCGTCTTCTGCTACTCCGGCCAGGGCAGCCATCCCGCGATGGGC 1850  
E L V F V Y S G Q G T Q H P A M G  
GAGCAGCTAGCCGCGCGCTTCCCGCTCTTCGCGCGGATCCATCAGCAGGT 1900  
5 E Q L A A A F P V F A E I H Q Q V  
GTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACGAGACCGGTTACG 1950  
W D L L D V P D L E V N E T G Y  
CCCAGCCGGCCCTGTTTCGCAATGCAGGTGGCTCTGTTCCGGGCTGCTGGAA 2000  
A Q P A L F A M Q V A L F G L L E  
10 TCGTGCGGCTGTACGACCGGACGCGGTGATCGGCCATTCGGTGGGTGAGCT 2050  
S W G V R P D A V I G H S V S E L  
TGCGGCTCGCTATGTGTCCGGGTGTGGTCTGTGGAGGATGCCTGCACCT 2100  
A A A Y V S G V W S L E D A C T  
15 TGGTGTGCGCGCGGGCTCCTCTGATGCAGGCTCTTCCCGCGGCTGGGGTG 2150  
L V S A R A R L M Q A L F A S G V  
ATGGTCTGCTGCTCCCGGTCTCGGAGGATGAGGCCCGGCGCGGTGCTGGGTGA 2200  
M V A V P V S E D E A R A V L G E  
GGGTGTGAGATCGCCGCGGTCACCGGCCGTCGTGCGGTGGTTCTCTCCG 2250  
G V E I A A V N G P S S V V L S  
20 GTGATGAGGCCCGCGGTGCTGCAGGCCGCGGAGGGGCTGGGGAAGTGGACG 2300  
G D E A A V L Q A A E G L G K W T  
CGGCTGCGGACCGACCGCGGTTCCATTCCGCCCCGATGGAACCCATGCT 2350  
R L A T S H A F H S A R M E F M L  
25 GGAGGAGTTCGCGCGGTGCGCGAAGGCCTGACCTACCGGACCGCGCAGG 2400  
E E F R A V A E G L T Y R T P Q  
TCTCCATGGCCGTTGGTGTATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450  
V S M A V G D Q V T T A E Y W V R  
CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500  
Q V R D T V R F G E Q V A S Y E D  
30 CGCCGTGTTTCGTGAGCTGGGTGCCGACCGGTCACTGGCCCGCCTGGTTCG 2550  
A V F V E L G A D R S L A R L V  
ACGGTGTGCGGATGCTGCACGGCGACACGAAATCCAGGCCGCGATCGGC 2600  
D G V A M L H G D H E I Q A A I G  
GCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCCACTGGCCCGCGCT 2650  
35 A L A H L Y V N G V T V D W F A L  
CCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCT 2700  
L G D A P A T R V L D L P T Y A  
TCCAGCACCAGCGCTACTGGCTCGAGTCCGGCACGCCCGCGCATCCGAC 2750  
F Q H Q R Y W L E S A R P A A S D  
40 GCGGGCCACCCCGTGTGGGTCCGGTATCGCCCTCGCCGGGTGCGCGGG 2800  
A G H P V L G S G I A L A G S P G  
CCGGGTGTTACCGGTTCCGTGCCGACCGGTGCGGACCGCGCGGTGTTTCG 2850  
R V F T G S V P T G A C R A V F  
TCGCCGAGCTGGCGCTGGCCCGCGGACGCGGTCCACTGCGCCACGGTC 2900  
45 V A E L A L A A A D A V D C A T V  
GAGCGGCTCGACATCGCCTCCGTGCCCGGCCGCGCGGCGCATGGCCGGAC 2950  
E R L D I A S V P G R P G H G R T  
GACCGTACAGACCTGGGTGACGAGCCGCGGACGACGCGCGCGCGCGGT 3000  
T V Q T W V D E P A D C G R R R  
50 TCACCGTGCACACCCGACCGCGGACGCCCCGTGGACGCTGCACGCGGAG 3050  
F T V H T R T G D A P W T L H A E  
GGGGTGTGCGCCCCCATGACCGGCCCTGCCCGATGCGGCGGACGCGGA 3100  
G V L R F H G T A L P D A A C A E  
GTGGCCCCACCGCGCGCGGTGCCCGCGGACGGGTGCGCGGTGTGTGGC 3150  
55 W P P P G A V P A D G L P G V W  
GCCGGGGGGACAGGTCTTCGCGGAGGCGGAGGTGGACGGACCGGACGGT 3200  
R R G D Q V F A E A E V C G F D G  
TTCGTGCTGACCCCGACCTGCTGACGCGGTCTTCTCCGCGGTGCGCGA 3250  
F V V H P D L L D A V F S A V G D  
60 CGGAAGCCGCCAGCGCGCGGATGGCGGACCTGACGGTGCACGCGTCCG 3300

SUBSTITUTE SHEET (RULE 26)

E S R Q P A G W R D L T V H A S  
 ACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCGACGGAGCCATG 3350  
 D A T V L R A C L T R R T D G A M  
 5 GATTTCGCGCCTTCGACGGCGCGCGCCTGCGCGTACTCACCCGGAGGC 3400  
 G F A A F D G A G L P V L T A E A  
 GGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTCCGACG 3450  
 V T L R E V A S P S G S E E S D  
 GCGTCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACGACGGT 3500  
 G L H R L E W L A V A E A V Y D G  
 10 GACCTGCCCCGAGGGACATGTCCTGATCACCGCGCGCCACCCCGACGACCC 3550  
 D L P E G H V L I T A A H P D D P  
 CGAGGACATACCCACCCGCGCCACACCCGCGCCACCCGCGTCTTGACCG 3600  
 E D I P T R A H T R A T R V L T  
 CGCTGCAACACCACCTCACCAACCGACACACCTCATCGTCCACACC 3650  
 15 A L Q H H L T T T D H T L I V H T  
 ACCACCGACCCCGCGCGCGCCACCGTCACCGGCTCACCCGACCCGCCA 3700  
 T T D P A G A T V T G L T R T A Q  
 GAACGAACACCCCAACCGCATCCGCCTCATCGAAACCGACACCCCCACA 3750  
 N E H P H R I R L I E T D H P H  
 20 CCGCCCTCCCGCTGCGCACTCGCCACCCCTCGACACCCCCACCTCCGC 3800  
 T P L P L A Q L A T L D H P H L R  
 CTCACCCACCAACCCCTCCACCAACCCCACTCACCCCCCTCCACACCAC 3850  
 L T H H T L H H P H L T P L H T T  
 CACCCACCCACCAACCCCCCTCAACCCCGAACACGCCATCATCATCA 3900  
 25 T P P T T T P L N P E H A I I I  
 CCGGCGGCTCCGGCACCCCTCGCGGCGATCCTCGCCCGCCACCTGAACCAC 3950  
 T G S G T L A G I L A R H L N H  
 CCCCACACCTACCTCCTCTCCCGCACCCCAACCCCGACGCCACCCCGG 4000  
 P H T Y L L S R T P P P D A T P G  
 30 CACCCACCTCCCGTGGGACGTGGGCGACCCCAACCACTCGCCACCAACC 4050  
 T H L P C D V G D P H Q L A T T  
 TCACCCACATCCCCCAACCCCTCACCGCCATCTTCCACACCGCGGCCACC 4100  
 L T H I P Q P L T A I F H T A A T  
 35 CTCGACGACGGCATCCTCCACGCCCTCACCCCGACCGCCTCACCACCGT 4150  
 L D D G I L H A L T P D R L T T V  
 CCTCCACCCCAAGGCAACGCGCGCTGGCACCTGCACCACCTCACCCAAA 4200  
 L H P K A N A A W H L H H L T Q  
 ACCAACCCCTCACCACTTCGTCTCTACTCCAGCGCGCGCGCGTCTCTC 4250  
 N Q P L T H F V L Y S S A A A V L  
 40 GCGAGCCCGGACAAGGAACTACGCGCGCGCAACGCCTTCCTCGACGC 4300  
 G S P Q G N Y A A A N A F L D A  
 CCTCGCCACCCACCGCCACACCCCTCGGCCAACCGCCACCTCCATCGCCT 4350  
 L A T H R H T L G Q P A T S I A  
 45 GGGGCATGTGGCACACCACGACCCCTCACCGGACAACCTCGACGACGCC 4400  
 W G M W H T T S T L T G Q L D D A  
 GACCGGGACCGCATCCGCGCGGCGGTTTCCTCCCGATCACGGACGACGA 4450  
 D R D I R R G G F L P I T D D E  
 GGGCATGGGGATGCAT  
 50 G

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

55 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50  
 Q L A E A L L T L V R E S T  
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCGCGACGGCGGC 100

A A V L G H V G G E D I P A T A A  
 GTTCAAGGACCTCGGCATCGACTCGCTCAGCCCGGTCCAGCTGCGCAACG 150  
 F K D L G I D S L T A V Q L R N  
 5 CTTTACCGAGCGGACCGGTGTGCGGTGAACCGCCACGGCGGTCTTTCGAC 200  
 A L T E A T G V R L N A T A V F D  
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACCTGACCGG 250  
 F P T P H V L A G K L G D E L T G  
 CACCCGCGCGCCCGTCTGTGCCCCGGACCGCGGCCACGGCCGGTGCACG 300  
 T R A P V V P R T A A T A G A H  
 10 ACGAGCCGCTGCGGATCGTGGGAATGGCCTGCCGCTGCCCGCGGGGTC 350  
 D E P L A I V G M A C R L P G G V  
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400  
 A S P E E L W H L V A S G T D A I  
 CACGGAGTTCCCGACGGACCGCGGTGGGACGTGACGCGATCTACGACC 450  
 15 T E F P T D R G W D V D A I Y D  
 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500  
 P D P D A I G K T F V R H G G F L  
 ACCGCGCGACAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550  
 T G A T G F D A A F F G I S P R E  
 20 GGGCCCTCGGATGGACCCCGCAGCAGCGGGTGCTCCTGGAGACGTGCTGGG 600  
 A L A M D P Q Q R V L L E T S W  
 AGSCGTTTCAAAGCCCGGCGATCAGCCCGGACTCGACCCCGCGGACGCGAC 650  
 E A F E S A G I T P D S T R G S D  
 ACCGCGGTGTTCTGTCGGCGCCTTCTCTACGTTACGGCACCGGTGCGGA 700  
 25 T G V F V G A F S Y G Y G T G A D  
 CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750  
 T D G F G A T G S Q T S V L S G  
 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800  
 R L S Y F Y G L E G P A V T V D T  
 30 GCGTGTCTGTCGCTCGCTGGTGGCGCTGCACCAGGCCGGGCGAGTCTGCTGCG 850  
 A C S S L V A L H Q A G Q S L R  
 CTCCGCGCAATGCTCGCTCGCCCTGGTCCGGCGGCGTACGGTGATGGCGT 900  
 S G E C S L A L V G G V T V M A  
 CTCCCGCGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950  
 35 S P G G F V E F S R Q R G L A P D  
 GGCCGGGCGAAGGCGTTCGGCGCGGGTGCAGGACGGCACGAGCTTCGCCGA 1000  
 G R A K A F G A G A D G T S F A E  
 GGGTGCCGTTGTGCTGATCGTCTGAGAGGCTCTCCGACGCCGAACGCAACG 1050  
 G A G V L I V E R L S D A E R N  
 40 GTCAACCGCTCTGGCGGTCTCGTGGTTCGGCGGTCAACCAGGATGGT 1100  
 G H T V L A V V R G S A V N Q D G  
 GCCTCCAACGGGCTGTGCGCGCCGAACGGGCCGTGCGAGGAGCGGGTGAT 1150  
 A S N G L S A P N G P S Q E R V I  
 CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGCGACGTGGACGCCG 1200  
 45 R Q A L A N A G L T P A D V D A  
 TCGAGGCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250  
 V E A H G T G T R L G D P I E A Q  
 GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCTGCTGCTGGG 1300  
 A V L A T Y G Q E R A T P L L L G  
 50 CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCCGCCG 1350  
 S L K S N I G H A Q A A S G V A  
 GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400  
 G I I K M V Q A L R H G E L P P T  
 CTGCACGCCGACGAGCCGTCGCCGCACGTGCACTGGACGGCCGGCGCCGT 1450  
 55 L H A D E P S P H V D W T A G A V  
 CGAAGTGTGACGTCCGGCCCGCGCGTGGCCCGAGACCGACCGGCCACGGC 1500  
 E L L T S A P P W P E T D P P R  
 GTCCCGCGGTCTCTGCTGCGGGGTGAGCGGCACCAACGCCCAAGTCTATC 1550  
 R A A V S S F G V S G T N A H V I  
 60 CTCGAGGCCGACCGGTAAACGGAGACGCCCGCGGCATCGCCTTCGGGTGA 1600

SUBSTITUTE SHEET (RULE 26)





Phage KC515 DNA was prepared using the procedure described in Genetic Manipulation of *Streptomyces*. A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of confluent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *Bam*HI and *Pst*I, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes *Bg*II and *Nsi*I and ligated into the compatible *Bam*HI and *Pst*I sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of *Streptomyces lividans* TK24 using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual edited by D. Hopwood *et al.* and overlaid with TK24 spores. After 16-24 hr, the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood *et al.*, *supra*). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya *et al.* (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

*Streptomyces hygroscopicus* ATCC 14891 (sec US Patent No. 3,244,592, issued 5 Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage ( $1 \times 10^8$  of each), and incubating on R2YE agar (Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D. Hopwood *et al.*) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 µg/ml) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by



replica plating onto thiostrepton containing agar medium. The PCR was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS/AT junction or the AT/DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains, followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

### Example 2

#### Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference; *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894,366, incorporated herein by reference; *S. sp.* MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S. sp.* MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem.* 256: 528-534, and Motamedi *et al.*, 1997, "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506," *Eur. J. Biochem.* 244: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces sp.* MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant gene clusters of the present invention differ from the naturally occurring gene clusters in that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

The naturally occurring module 8 sequence for the MA6548 strain is shown below, followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

```

30      GCATCGGCTGTACGAGGCGGCTACCGCGCACCGGAAGTCCCGTGGTGGTG 50
          M R L Y E A A R R T G S P V V V
          GCGGCCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
35      A A A L D D A P D V P L L R G L R

```

CCGTACGACCGTCCGGCGGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
R T T V R R A A V R E R S L A D  
GCTCGCGGTGCTGCGCGACGAGCGCGCGGACGCGCTCCCTCGCGTTCCG 200  
E S P T C P T T S A P T P P S R S  
5 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCGGGAAGACAT 250  
S W N S T A T V L G H L G A E D I  
CCCGGCGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACC CGG 300  
P A T T T F K E L G I D S L T A  
TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
10 V Q L R N A L T T A T G V R L N A  
ACAGCGGTCTTCGACTTTCCGACGCCGCGCGCTCGCCGCGAGACTCGG 400  
T A V F D F P T P R A L A A R L G  
CGACGAGCTGGCCGGTACCCGCGCGCCCGTCCGCGCCCGGACCGCGGCCA 450  
D E L A G T R A P V A R T A A  
15 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
T A A A H D E P L A I V G M A C R  
CTGCCGGGGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
L P G G V A S P Q E L W R L V A S  
CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600  
20 G T D A I T E F P A D R G W D V  
ACGCGCTCTACGACCCGACCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
C A L Y D P D P D A I G K T F V R  
CACGGCGGCTTCTCGACGGTGGACCGGCTTCGACGCGGCGTTCTTCGG 700  
H G G F L D G A T G F D A A F F G  
25 GATCAGCCCCGCGCGAGGCCCTGGCCATGGACCCGACGAACGGGTGCTCC 750  
I S P R E A L A M D P Q Q R V L  
TGGAGACGTCTGGGAGGCGTTCGAAAGCGCGGCATCACCCCGGACGCG 800  
L E T S W E A F E S A G I T P D A  
GCGCGGGGCGAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850  
30 A R G S D T G V F I G A F S Y G Y  
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGACAGCA 900  
G T G A D T N G F G A T G S Q T  
GCGTGCTCTCCGGCCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
S V L S G R L S Y F Y G L E G P S  
35 GTCACGGTTCGACACCGCTGTCGTCGTCGCTCGCCCTGCACCAGGC 1000  
V T V D T A C S S S L V A L H Q A  
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050  
G Q S L R S G E C S L A L V G G  
TCACGGTGATGGCGTCCGCCGCGGATTCGTGAGTTCTCCCGGCAGCGC 1100  
40 V T V M A S P G G F V E F S R Q R  
GGGCTCGCGCGGACGGGCGGCGGAAGGCGTTCGCGCGGGGCGCGGACGG 1150  
G L A P C G R A K A F G A G A D G  
TACGAGCTTCGCGGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200  
T S F A E G A G A L V V E R L S  
45 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250  
D A E R H G H T V L A L V R G S A  
GCTAACTCCGACGGCGCGTTCGAACGGTCTGTGCGGCGCGGAACGGCCCTC 1300  
A N S D G A S N G L S A P N G P S  
CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCGG 1350  
50 Q E R V I H Q A L A N A K L T P  
CCGATGTTCGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D R A T  
55 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCGG 1500  
P L L L G S L K S N I G H A Q A  
TCTCAGGGGTTCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
GAACTGCCGCGGACACTCCACGCGGACGAGCCGTGCGCCGACGTGACTG 1600  
60 E L P P T L H A D E P S P H V D W

SUBSTITUTE SHEET (RULE 26)

GACGCGCGGTGCCCTCGAGCTCCTGACGTCGGCCCGGCCCTGGCCGGGGA 1650  
T A G A V E L L T S A R P W P G  
CGGGTGGCCCGCGCCGCGCTGCCGTCTCGTCTCGGCGTGGAGCGGCACG 1700  
T G R F R R A A V S S F S G T  
5 AACGCGCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750  
N A H I I L E A G P V K T G P V E  
GGCAGGAGCGATCGAGGCAGGACCGGTCTGAAGTAGGACCGGTCTGAGGCTG 1800  
A S A I E A G P V E V G P V E A  
10 GACCGCTCCCGCGCGCCGCGCTCAGCACCGGCGGAGACTTTCCGCTG 1850  
G P L P A A P P S A P S E D L P L  
CTCGTCTCGGCGCGCTTCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
L V S A R S P E A L D E Q I G R L  
GCGCGCCTATCTCGACACCGGCGCGGCGTCTGACCGGCGGCGCTGGCGC 1950  
R A Y L D T G P G V D R A A V A  
15 AGACACTGGCCCGCGCTACGCACTTCACCCACCGGCGCGTACTGCTCGGG 2000  
Q T L A R R T H F T H R A V L L G  
GACACCGTTCATCGGCGCTCCCCCGCGGACCGGCGGACGAACTCGTCTT 2050  
D T V I G A P P A D Q A D E L V F  
CGTCTACTCCGGTCAAGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100  
20 V Y S G Q G T Q H P A M G E Q L  
CGGCGCGCTTCCCGGTGTTCCCGGATGCCTGGCAGGACGCGCTCCGACGG 2150  
A A A F P V F A D A W H D A L R R  
CTCGACGACCCCGACCCGACGACCCACACGGAGCCAGCACACGCTCTT 2200  
L D D P D P H D P T R S Q H T L F  
25 CGCCACCGGCGCGCTTACCGCCCTCCTGAGGTCTCTGGGACATCACGC 2250  
A H Q A A F T A L L R S W D I T  
CGCACGCGCTCATCGGCCACTCGCTCGGCGAGATCACCGCGCGTACGCC 2300  
P H A V I G H S L G E I T A A Y A  
30 GCGGGATCCTGTCTGCTCGACGACGCTGCACCTGATCACACGCGTGC 2350  
A G I L S L D D A C T L I T T R A  
CCGCCTCATGCACACGCTTCCGCGCGCGGCGCCATGGTACCGTGTCTGA 2400  
R L M H T L P P P G A M V T V L  
CCAGCGAGGAGGAGGCGCGTCAAGCGCTGCGGCGGGCGTGGAGATCGCC 2450  
T S E E E A R Q A L R P G V E I A  
35 GCGGTCTTCCGCGCGCACTCCGTCTGTCTCTCGGGCGACGAGGACGCGGT 2500  
A V F G P H S V V L S G D E D A V  
GCTCGACGTCTGCACAGCGGCTCGGCATCCACCACCGTCTGCCGCGCGCGC 2550  
L D V A Q R L G I H R L P A P  
ACGCGGGCACTCCGCGCATGGAACCGTGGCGCGGAGCTGCTCGCC 2600  
40 H A G H S A H M E P V A A E L L A  
ACCACTCGCGAGCTCCGTTACGACCGGCCCCACACCGCCATCCCGAACGA 2650  
T T R E L R Y D R P H T A I P N D  
CCCCACCAACCGCGAGTACTGGGCGGAGCAGGTCCGCAACCCCGTGTGT 2700  
P T T A E Y W A E Q V R N P V L  
45 TCCACGCCCCACCCAGCGGTACCCCGACGCGGTCTTCTGTCAGATCGGC 2750  
F H A H T Q R Y P D A V F V E I G  
CCCGGCCAGGACCTCTCACCGCTGGTCTGACGGCATCGCCCTGAGAACGG 2800  
P G Q D L S P L V D G I A L Q N G  
50 CACGGCGGACGAGGTGCACGCGCTGCACACCGCGCTCGCCCGCTCTTCA 2850  
T A D E V H A L H T A L A R L F  
CACGCGGCGCCACGCTCGACTGGTCCCGCATCCTCGGCGGTGCTTCGCGG 2900  
T R G A T L D W S R I L G G A S R  
CACGACCTGACGTCCCTCTGACGCTTCCAGCGGCGTCCCTACTGGAT 2950  
H D P D V P S Y A F Q R R P Y W I  
55 CGAGTCCGCTCCCCCGSCACGGCCGACTCGGGCCACCCCGTCTCTCGGCA 3000  
E S A P P A T A D S G H P V L G  
CCGGAGTCCCGCTCCCGGGTCCCGGGCGGGGTGTTACGCGGTCCCGTG 3050  
T G V A V A G S P G R V F I G P V  
CCCGCGGTGCGGACCGCGCGGTGTTTCATCGCCGAACTGGCGCTCGCCGC 3100  
60 P A G A D R A V F I A E L A L A A

SUBSTITUTE SHEET (RULE 26)

CGCCGACGCCACCGACTGCGCCACGGTCTGAACAGCTCGACGTCACCTCCG 3150  
A D A T D C A T V E Q L D V T S  
TGCCCGGCGGATCCGCCCGCGGCAGGGCCACCGCGCAGACCTGGGTGAT 3200  
P G G S A R G R A T A T W V D  
5 GAACCCGCGCGCGACGGGCGGCGCGCTTCACCGTCCACACCCGCGTCGG 3250  
E P A A D G R R R F T V H T R V G  
CGACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGGCGCG 3300  
D A P W T L H A E G V L R P G R  
TGCCCCAGCCGAAGCCGTCGACACCGCCTGGCCCCCGCGGGCGCGGTG 3350  
10 V P Q P E A V D T A W P P P G A V  
CCCGCGGACGGGCTGCCCGGGGCGTGGCGACGCGCGGACCAGGTCTTCGT 3400  
P A D G L P G A W R R A D Q V F V  
CGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGC 3450  
E A E V D S P D G F V A H P D L  
15 TCGACGCGGTCTTCTCCGCGGTCCGGCAGGGAGCCGCCAGCCGACCGGA 3500  
L D A V F S A V G D G S R Q P T G  
TGGCGCGACCTCGCGGTGCACGCGTCGGACGCCACCGTGTGCGCGCCTG 3550  
W R D L A V H A S D A T V L R A C  
CCTCACCCGCGCGACAGTGGTGTCTGAGCTCGCCGCCTTCGACGGTG 3600  
20 L T R R D S G V V E L A A F D G  
CCGGAATGCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTCCGG 3650  
A G M P V L T A E S V T L G E V A  
TCGGCAGGCGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTT 3700  
S A G G S D E S D G L L R L E W L  
25 GCCGGTGGCGGAGGGCCACTACGACGGTGCCGACGAGTCCCCGAGGGCT 3750  
P V A E A H Y D G A D E L P E G  
ACACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCCAAC 3800  
Y T L I T A T H P D D P D P T N  
CCCCACAACACACCCACACGCACCCACACACAAACCACACGCGTCTCAC 3850  
30 P H N T P T R T H T Q T T R V L T  
CGCCCTCCAACACCACCTCATCACCACCAACCACACCCTCATCGTCCACA 3900  
A L Q H H L I T T N H T L I V H  
CCACCACCGACCCCCCAGGCGCCGCGCTCACCGGCCTCACCCGACCCGCA 3950  
T T T D P P G A A V T G L T R T A  
35 CAAAACGAACACCCCGCGGCTCATCAGAAACCCACCCCAAC 4000  
Q N E H P G R I H L I E T H H P H  
CACCCCACTCCCCCTCACCAACTCACACCCTCCACCAACCCCACTAC 4050  
T P L P L T Q L T T L H Q P H L  
GCCTCACCAACAACACCTCCACACCCCCACCTCACCCCATCACCAAC 4100  
40 R L T N N T L H T P H L T P I T T  
CACCACAACACCACCAACACCCCAACACCCCAACCCCTCAACCCCAA 4150  
H H N T T T T P N T P P L N P N  
CCACGCCATCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGGCATCCTCG 4200  
H A I L I T G G S G T L A G I L  
45 CCGGCCACCTCAACCACCCCAACACCTACCTCCTCTCCGCGACACCACCA 4250  
A R H L N H P H T Y L L S R T P P  
CCCCCACCACACCCGGCACCCACATCCCTGCGACCTCACCGACCCAC 4300  
P P T T P G T H I P C D L T D P T  
CCAAATCACCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCT 4350  
50 Q I T Q A L T H I P Q P L T G I  
TCCACACCGCGCCACCCCTCGACGACGCCACCCCTCACCAACCTCACCC 4400  
F H T A A T L D D A T L T N L T P  
CAACACCTCACCAACCCCTCAACCCAAAGCCGACGCCGCTGGCACCT 4450  
Q H L T T T L Q P K A D A A W H L  
55 CCACCACCACCCCAACCAACCCCTCACCACTTCGTCTCTACTCCA 4500  
H H H T Q N Q P L T H F V L Y S  
GCGCGCGCGCACCCCTCGGCAGCCCCGGGCAAGCCAACTACGCCGCGCC 4550  
S A A A T L G S P G Q A N Y A A A  
AACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCAAGGACAACC 4600  
60 N A F L D A L A T H R H T Q G Q P

CGCCACCACCATCGCCTGGGGCATGTGGCACACCACCACACTCACCA 4650  
 A T T I A W G M W H T T T T L T  
 GCCAACTCACCGACAGCGACCGCGACCGCATCCGCGCGGGCGGCTTCCTG 4700  
 S Q L T D S D R D R I R R G G F L  
 5 CCGATCTCGGACGACGAGGGCATGC  
 P I S D D E G M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

10 GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 M R L Y E A A R R T G S P V V V  
 GCGGCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 A A A L D D A P D V P L L R G L R  
 GCGTACGACCGTCCGGCGTGGCGCGTCCGGGAACGCTCTCTCGCCGACC 150  
 15 R T T V R R A A V R E R S L A D  
 GCTCGCGTGTGCCGACGACGAGCGCGCGGACGCCTCCCTCGCGTTTCG 200  
 R S P C C P T T S A P T P P S R S  
 TCCTGGAACAGCACCGCCACCGTGTCTCGGCCACCTGGGCGCGGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 20 CCGGCGGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300  
 P A T T T F K E L G I D S L T A  
 TCCAGCTGCGCAACGCGCTGACCAACGGCGACCGCGTACGCGCTCAACGCC 350  
 V Q L R N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCGACCGCGCGCGCTCGCGCGAGACTCGG 400  
 25 T A V F D F P T P R A L A A R L G  
 CGACGAGCTGGCCGGTACCCGCGCGCCCGTTCGCGGCGCGGACCGCGGCCA 450  
 D E L A G T R A P V A A R T A A  
 CCGCGGCGCGGACGACGACGACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 30 CTGCCGCGGGGGTTCGCGTCCGACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 L P G G V A S P Q E L W R L V A S  
 CGGCACCGACGCCATCACGGAGTTCGCCGCGGACCGCGGCTGGGACGTGG 600  
 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 35 D A L Y D P D A I G K T F V R  
 CACGGCGGCTTCCTCGACGGTGGACCGGCTTCGACGCGGCGTTCTTCGG 700  
 H G G F L D G A T G F D A A F F G  
 GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q Q R V L  
 40 TGGAGACGTCTGGGAGGCGTTTCAAAGCGCGGGCATCACCCCGGACGCG 800  
 L E T S W E A F E S A G I T P D A  
 GCGCGGGGACGACACCGCGGTTCATCGGCGCGTTCCTACGGGTA 850  
 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGCCA 900  
 45 G T G A D T N G F G A T G S Q T  
 GCGTGTCTCTCGGCGCGCTCTCTGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 S V L S G R L S Y F Y G L E G P S  
 GTCACGGTTCGACACCGCCTGCTCGTCTGTCAGTGGTCCGCTGCACAGGC 1000  
 V T V D T A C S S S L V A L H Q A  
 50 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050  
 G Q S L R S G E C S L A L V G G  
 TCACGGTGATGGCGTTCGCGCGGCGGATTCGTGAGTTCCTCCGCGACGCG 1100  
 V T V M A S P G G F V E F S R Q R  
 GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150  
 55 G L A P D G R A K A F G A G A D G  
 TACGAGCTTCGCGAGGGCGCGCTTCGCGCGGCTTCGAGCGGCTCTCCG 1200  
 T S F A E G A G A L V V E R L S  
 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250  
 D A E R H G H T V L A L V R G S A

GCTAACTCCGACGGCGCGTCGAACGGTCTGTGCGGCGCGGAACGGCCCCCTC 1300  
A N S D G A S N G L S A P N G P S  
CCAGGAACCGGTCATCCACCAGGCCCTCGCGAACCGGAACTCACCCTCCG 1350  
Q E R V I H Q A L A N A K L T P  
5 CCGATGTGACGCGCGTCGAGGCGCACGGCACCGGCACCGCGCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D P A T  
10 GCCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCGG 1500  
P L L L G S L K S N I G H A Q A  
CGTCAGGGGTGCGCGGATCATCAAGATGGTGACGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
GAACTGCCGCGGACACTGCACGCGGACGAGCCCTCGCGCACGTCGACTG 1600  
E L P P T L H A D E P S P H V D W  
15 GACGGCCGGTGCCGTCGAGCTCCTGACGTGCGGCCCGCGCTGGCCGGGGA 1650  
T A G A V E L L T S A R P W P G  
CCGGTCGCCCTAGGCGGGCAGGCGTGTGCTCCTTCGGGATCAGTGGCACC 1700  
T G R P R R A G V S S F G I S G T  
AAGCCCCACGTCTCTGGAAGCGCACCCCCCACTCAGCCTGCGGACAA 1750  
N A H V I L E S A P P T Q P A D N  
20 CGCGGTGATCGAGCGGGCACCAGGAGTGGGTGCGGTTGGTGATTTCGGCCA 1800  
A V I E R A P E W V P L V I S A  
GGACCCAGTCGGCTTTGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTG 1850  
R T Q S A L T E H E G R L R A Y L  
25 GCGGCGTCGCCCGGGGTGGATATGCGGGCTGTGGCATCGACGCTGGCGAT 1900  
A A S P G V D M R A V A S T L A M  
GACACGGTCGGTGTTTCGAGCACCGTGCCGTGCTGCTGGGAGATGACACCG 1950  
T R S V F E H R A V L L G D D T  
30 TCACCGGCACCGCTGTGTCTGACCCTCGGGCGGTGTTGCTCTTCCCGGGA 2000  
V T G T A V S D P R A V F V F P G  
CAGGGGTGCGCAGCGTGTGGCATGGGTGAGGAACTGGCCGCCGCGTTCCC 2050  
Q G S Q R A G M G E E L A A A F P  
CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCG 2100  
V F A I H Q Q V W D L L D V P  
35 ATCTGGAGGTGAACGAGACCGGTTACGCCAGCCGGCCCTGTTTCGCAATG 2150  
D L E V N E T G Y A Q P A L F A M  
CAGGTGGCTCTGTTTCGGGCTGCTGGAATCGTGGGGTGTACGACCGGACGC 2200  
Q V A L F G L L E S W G V R P D A  
GGTGATCGGCCATTCGGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGG 2250  
40 V I G H S V G E L A A A Y V S G  
TGTGGTCTGTTGGAGGATGCGCTGCACTTTGGTGTGCGGCGGGCTCGTCTG 2300  
V W S L E D A C T L V S A R A R L  
ATGCAGGCTCTGCCCGCGGGTGGGGTGATGGTCTGCTGCCCGGTCTCGGA 2350  
M Q A L P A G G V M V A V P V S E  
45 GGATGAGGCCCGGGCCGTGCTGGGTGAGGGTGTGGAGATCGCCGCGGTCA 2400  
D E A R A V L G E G V E I A A V  
ACGGCCCGTCTGCGGTGGTTCTCTCCGGTGATGAGGCCCGCGTGTGCTG 2450  
N G P S S V V L S G D E A A V L Q  
GCCGCGGAGGGGTGGGGAAGTGACGCGGCTGGCGACCGAGCCACGCGTT 2500  
50 A A E G L G K W T R L A T S H A F  
CCATTCGCCCGTATGGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCG 2550  
H S A R M E P M L E E F R A V A  
AAGGCCTGACCTACCGGACCGCGCAGGTCTCCATGGCCGTTGGTGATCAG 2600  
E G L T Y R T P Q V S M A V G D Q  
55 GTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTT 2650  
V T T A E Y W V R Q V R D T V R F  
CGGCGAGCAGGTGGCCTCGTACGAGGACCGCGTGTTCGTCGAGCTGGGTG 2700  
G E Q V A S Y E D A V F V E L G  
CCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTGCGGATGCTGCACGGC 2750  
60 A D R S L A R L V D G V A M L H G

SUBSTITUTE SHEET (RULE 26)

GACCACGAAATCCAGGCCGCGATCGGGCGCCCTGCCCCACCTGTATGTCAA 2800  
D H E I Q A A I G A L A H L Y V N  
CGGGGTACGGTCCGACTGGCCCCGCGCTCCTGGGCGATGCTCCGGCAACAC 2850  
G T T V D W P A L L G D A P A T  
5 GGGTGTGACCTTCCGACATACGCCTTCCAGCACCAGCGCTACTGGCTC 2900  
R V L D L P T Y A F Q H Q R Y W L  
GAGTCGGCTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCAC 2950  
E S A P P A T A D S G H P V L G T  
10 CGGASTCGCGCTCCCCGGGTGCGCGGGCCGGGTGTTACGGGTCCCGTGC 3000  
G V A V A G S P G R V F T G P V  
CCCCCGGTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCC 3050  
P A G A C R A V F I A E L A L A A  
GCGGACGCCACCGACTCGGCCACGGTCTGAACAGCTCGACGTCACCTCCGT 3100  
A D A T D C A T V E Q L D V T S V  
15 GCGCGGCGGATCCGCCCCGCGGACGGGCCACCGCGCAGACCTGGGTGATG 3150  
P G G S A R G R A T A Q T W V D  
AACCCCGCGCGGACGGGCGGCGCGCTTACCGTCCACACCCGCGTCCGC 3200  
E P A A D G R R R F T V H T R V G  
GACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGCGCGCT 3250  
20 D A P W T L H A E G V L R P G R V  
GCCCCAGCCCGAAGCCGTGACACCGCTGGCCCCCGCGGGCGCGGTGC 3300  
P Q P E A V D T A W P P P G A V  
CCGCGGACGGGCTGCCCCGGGCGTGGCGACGCGCGGACCAGGTCTTCGTC 3350  
P A D G L P G A W R R A D Q V F V  
25 GAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCT 3400  
E A E V D S P D G F V A H P D L L  
CGACGCGGTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGAT 3450  
D A V F S A V G D G S R Q P T G  
GGCGGACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGC 3500  
30 W R D L A V H A S D A T V L R A C  
CTCACCCGCGGACAGTGGTGTGCTGGAGCTCGCCGCTTCGACGGTGC 3550  
L T R R D S G V V E L A A F D G A  
CGGAATGCCGGTGTCTACCGCGGAGTGGTGACGCTGGGCGAGGTGCGCT 3600  
G M P V L T A E S V T L G E V A  
35 CGGCAGSCGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTG 3650  
S A G G S D E S D G L L R L E W L  
CCGGTGGCGGAGGCCCCACTACGACGGTGCCGACGAGCTGCCCCGAGGGCTA 3700  
P V A E A H Y D G A D E L P E G Y  
CACCTCATCACCGCCACACACCCCGACGACCCCGACGACCCCAACCAACC 3750  
40 T L I T A T H P D D P D D P T N  
CCCACAACACACCCACACGCACCCACACACAAACCACACGGTCTCTCACC 3800  
P H N T P T R T H T Q T T R V L T  
GCCCTCCAACACCACCTCATCACCACCAACCACACCTCATCGTCCACAC 3850  
A L Q H H L I T T N H T L I V H T  
45 CACCACGACCCCCAGGCGCGCGCTCACCGGCTCACCCGCACCGCAC 3900  
T T D P P G A A V T G L T R T A  
AAAACGAACACCCCGCGCGCATCCACCTCATCGAAACCCACACCCCCAC 3950  
Q N E H P G R I H L I E T H H P H  
ACCCCACTCCCCCTCACCCAACTCACACCCCTCCACCAACCCCACTACG 4000  
50 T P L P L T Q L T T L H Q P H L R  
CCTCACCAACAACACCTCCACACCCCCACCTCACCCCATCACCAACC 4050  
E T N N T L H T P H L T P I T T  
ACCACAACACCAACCAACCAACCCCAACACCCCAACCCCTCAACCCCAAC 4100  
H H N T T T T T P N T P P L N P N  
55 CACGCCATCCTCATCACCGCGGGTCCGGCACCCCTCGCCGGCATCCTCGC 4150  
H A I L I T G G S G T L A G I L A  
CGGCACTCAACCAACCCCAACACCTACCTCCTCTCCCGCACACCAACC 4200  
R H L N H P H T Y L L S R T P P  
CCCCCACCACCCCGGCACCCACATCCCTGCGACCTCACCGACCCCAACC 4250  
60 P P T T P G T H I P C D L T D P T

SUBSTITUTE SHEET (RULE 26)

CAAATCACCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCTT 4300  
 Q I T Q A L T H I P Q P L T G I F  
 CCACACCGCGCCACCCTCGACGACGCCACCCTCACCAACCTCACCCCCC 4350  
 H T A A T L D D A T L T N L T P  
 5 AACACCTCACCAACCCTCCAACCCAAAGCCGACGCCGCTGGCACCTC 4400  
 Q H L T T T L Q P K A D A A W H L  
 CACCACCACACCCAAAACCAACCCCTCACCACTTCGTCTCTACTCCAG 4450  
 H H H T Q N Q P L T H F V L Y S S  
 CGCCGCGCCACCCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCCGCCA 4500  
 10 A A A T L G S P G Q A N Y A A A  
 ACGCCTTCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCC 4550  
 N A F L D A L A T H R H T Q G Q P  
 GCCACCACCTCGCCTGGGGCATGTGGCACACCACCACCACTCACCAG 4600  
 A T T I A W G M W H T T T T L T S  
 15 CCAACTACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCCTGC 4650  
 Q L T D S D R D R I R R G G F L  
 CGATCTCGGACGACGAGGGCATGC  
 P I S D D E G M

20 The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module  
 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 M R L Y E A A R R T G S P V V V  
 GCGGCCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100  
 25 A A A L D D A P D V P L L R G L R  
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150  
 R T T V R R A A V R E R S L A D  
 GCTCGCCGTGCTGCCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTCTG 200  
 R S P C C P T T S A P T P P S R S  
 30 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
 S W N S T A T V L G H L G A E D I  
 CCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300  
 P A T T T F K E L G I D S L T A  
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
 35 V Q L R N A L T T A T G V R L N A  
 ACAGCGGTCTTCGACTTTCGACGCCGCGCGCGCTCGCCGCGAGACTCGG 400  
 T A V F D F P T P R A L A R L G  
 CGACGAGCTGCCCGGTACCCGCGCGCCGTCGCGGCCCGGACCGCGGCCA 450  
 D E L A G T R A P V A A R T A A  
 40 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500  
 T A A A H D E P L A I V G M A C R  
 CTGCCGGGCGGGGTCCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550  
 L P G G V A S P Q E L W R L V A S  
 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600  
 45 G T D A I T E F P A D R G W D V  
 ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
 D A L Y D P D P D A I G K T F V R  
 CACGGCGGCTTCCTCGACCGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700  
 H G G F L D G A T G F D A A F F G  
 50 GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750  
 I S P R E A L A M D P Q Q R V L  
 TGGAGACGTCTCGGAGGCGTTGGAAGCGCGGGCATCACCCCGGACGCG 800  
 L E T S W E A F E S A G I T P D A  
 GCGCGGGGCGAGACACCGCGGTGTTTCATCGGCGGTTCTCTACGGGTA 850  
 55 A R G S D T G V F I G A F S Y G Y  
 CGGCACGGGTGCGGATACCAAGGCTTCGGCGGACAGGGTTCGACAGCA 900  
 G T G A D T N G F G A T G S Q T  
 GCGTGCTCTCGGCGCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
 S V L S G R L S Y F Y G L E G P S



CTCACGGTTCGACACCGCCTGCTCGTCTCGTCACTGGTCCGCCCTGCACCAGGC 1000  
V T V D T A C S S S L V A L H Q A  
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050  
G Q S L R S G E C S L A L V G G  
5 TCACGGTGATGGCTCGCCCGGGCGGATTCTGTCGAGTTCTCCCGGCAGCGC 1100  
V T V M A S P G G F V E F S R Q R  
GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150  
G L A P D G R A K A F G A G A D G  
TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200  
10 T S F A E G A G A L V V E R L S  
ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250  
D A E R H G H T V L A L V R G S A  
GCTAACTCCGACGGCGCGTCAACCGGTCTGTCGGCGCCGAACGGCCCTC 1300  
A N S D G A S N G L S A P N G P S  
15 CCAGGAACCGGTATCCACCAGGCCCTCGCGAACCGGAACTCACCCCCG 1350  
Q E R V I H Q A L A N A K L T P  
CCGATGTCGACGCGGTTCGAGGCGCACGGCACCGGCCCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
20 P I E A Q A L L A T Y G Q D R A T  
GCCCTGCTGCTCGGCTCGCTGAAGTCAACATCGGGCACGCCAGGCCG 1500  
P L L L G S L K S N I G H A Q A  
CGTCAGGGGTTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
A S G V A G I I K M V Q A I R H G  
25 GAAGTCCCGGACACTGCACGCGGACGAGCCGTTCGCCGACGTCGACTG 1600  
E L P P T L H A D E P S P H V D W  
GACGGCCGGTTCGCTCGAGCTCCTGACGTCGGCCCGGCCGTGGCCGGGGA 1650  
T A G A V E L L T S A R P W P G  
CCGGTTCGCCCTAGGCGGGCGGGCGTGTCTCTTCGGAGTCAGCGGCACC 1700  
30 T G R P R R A G V S S F G V S G T  
AACGCCCCAGTTCCTTGGAGAGCGCACCCCCCGCTCAGCCCCGCGGAGGA 1750  
N A H V I L E S A P P A Q P A E E  
GGCGCAGCCTGTTGAGACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGG 1800  
A Q P V E T P V V A S D V L P L  
35 TGATATCGGCCAAGACCCAGCCCGCCCTGACCGAACACGAAGACCGGCTG 1850  
V I S A K T Q P A L T E H E D R L  
CGCGCCTACCTGGCGGCGTTCGCCGGGGCGGATATACGGGCTGTGGCATC 1900  
R A Y L A A S P G A D I R A V A S  
3ACCGTGGCGGTGACACGGTTCGTTTCGAGCACCGCGCGTACTCCTTG 1950  
40 T L A V T R S V F E H R A V L L  
GAGATGACACCGTCAACGGCACCGCGGTGACCGACCCAGGATCGTGTTC 2000  
G D D T V T G T A V T D P R I V F  
GTCTTTCCCGGGCAGGGGTGGCAGTGGCTGGGGATGGGCAGTGCAGTGGC 2050  
V F P G Q G W Q W L G M G S A L R  
45 CGATTCTGTCGGTGGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCGT 2100  
D S S V V F A E R M A E C A A A  
TGCGCGAGTTCTGTTGGACTGGGATCTGTTACGGTTCTGGATGATCCGGCG 2150  
L R E F V D W D L F T V L D D P A  
GTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGT 2200  
50 V V D R V D V V Q P A S W A M M V  
TTCCCTGGCCCGGGTGTGGCAGGCGGCCGGTGTGCGGCCGGATGCGGTGA 2250  
S L A V W Q A A G V R P D A V  
TCGGCCATTTCGAGGGTGAGATCGCCGACGCTTGTGTGGCGGGTSCGGTG 2300  
I G H S Q G E I A A A C V A G A V  
55 TCACTACCGGATGCCGCCCGGATCGTGACCTTGCGCAGCCAGGCGATCGC 2350  
S L R D A A R I V T L R S Q A I A  
CCGSGGCCTTCGCGGCCCGGGCGGATGGCATCCGTCCGCCCTGCCCGGCC 2400  
E G L A G R G A M A S V A L P A  
AGGATGTCGAGCTGGTTCGACGGGGCCTGGATCGCCGCCACAAACGGGGCC 2450  
60 L D V E L V D G A W I A A H N G P

GCCTCCACCGTGTATCGCGGGCACCCCGGAAGCGGTGACCATGTCTCAC 2500  
A S T V I A G T P E A V D H V L T  
CGCTCATGAGGCACAAGGGGTGCGGGTGGCGGGATCACCGTGGACTATG 2550  
A H E A G V R V R R I T V D Y  
5 CCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAACACTACTCGACATC 2600  
A S H T P H V E L I R D E L L D I  
ACTAGCGACAGCAGCTCCGAGACCCCGCTCGTGCCGTGGCTGTGACCGT 2650  
T S D S S S Q T P L V P W L S T V  
GGACGGCACCTGGSTCGACAGCCCGCTGGACGGGGAGTACTGGTACCGGA 2700  
10 D G T W V D S P L D G E Y W Y R  
ACCTGCGTGAACCGGTGCGTTTCCACCCCGCGCTCAGCCAGTTGCAGGCC 2750  
N L R E P V G F H P A V S Q L Q A  
CAGGCGCACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGCA 2800  
Q G D T V F V E V S A S P V L L Q  
15 GGCGATGGACGACGATGTCGTCACGGTTGCCACGCTGCGTCGTGACGACG 2850  
A M D D D V V T V A T L R R D D  
GCGACGCTACCCGATGCTCACC GCCCTGGCACAGGCCTATGTCCACGGC 2900  
G D A T R M L T A L A Q A Y V H G  
GTCACCGTGGACTGGCCCGCATCCTCGGCACCACCACAACCCGGGTACT 2950  
20 V T V D W P A I L G T T T T R V L  
GGACCTTCCGACCTACGCCCTTCCAACACCAGCGGTACTGGCTCGAGTCGG 3000  
D L P T Y A F Q H Q R Y W L E S  
CTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCACCGGAGTC 3050  
A P P A T A D S G H P V L G T G V  
25 GCGGTGCGCGGGTGGCGGGCGGGTGTTCACGGGTCCCGTGCCCGCCGG 3100  
A V A G S P G R V F T G P V P A G  
TGCGGACCGCGCGGTGTTCATCGCCGAACGTGGCGCTCGCCGCCGCCGACG 3150  
A D R A V F I A E L A L A A A D  
CCACCGACTGCGCCACGGTCAACAGCTCGACGTCACCTCCGTGCCCCGGC 3200  
30 A T D C A T V E Q L D V T S V P G  
GGATCCGCCCCGCGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCGC 3250  
G S A R G R A T A Q T W V D E P A  
CGCCGACGGGGCGCGCGCTTACCGTCCACACCCCGCTCGGCGACGCC 3300  
A D G R R R F T V H T R V G D A  
35 CGTGGAGCTGCACGCGGAGGGGTCTCCGCCCGCGCGCTGCCCCAG 3350  
P W T L H A E G V L R P G R V P Q  
CCCGAAGCGGTGACACCGCTGGCCCCCGCGGGCGCGGTGCCCCGCGGA 3400  
P E A V D T A W P P P G A V P A D  
CGGGGTGCCCCGGGGCGTGGCGACGCGCGGACAGGTCTTCGTGGAAGCCG 3450  
40 G L P G A W R R A D Q V F V E A  
AAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGCG 3500  
E V D S P D G F V A H P D L L D A  
GTCTTCTCCGCGGTGCGGACGCGGAGCCGCCAGCCGACCGGATGGCGCGA 3550  
V F S A V G D G S R Q P T G W R D  
45 CCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGCCTCACCC 3600  
L A V H A S D A T V L R A C L T  
GCCGCGACAGTGGTGTGCTGGAGCTCGCCGCTTCGACGGTGCCGGAATG 3650  
R R D S G V V E L A A F D G A G M  
CCGGTGTCTACCGCGGAGTGGTGACGCTGGGCGAGGTGCGGTGCGCAGG 3700  
50 P V L T A E S V T L G E V A S A G  
CGGATCCGACGAGTGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTGG 3750  
G S D E S D G L L R L E W L P V  
CGGAGGCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTACACCCCTC 3800  
A E A H Y D G A D E L P E G Y T L  
55 ATCACCGCCACACACCCCGACGACCCCGACGACCCCAACACCCCAACAA 3850  
I T A T H P D D P D D P T N P H N  
CACACCCACACGACCCACACACAAACACACGCGTCTCACCGCTCTCC 3900  
T P T R T H T Q T T R V L T A L  
AACACCACCTCATCACCAACACACCCCTCATCGTCCACACCAACCACC 3950  
60 Q H H L I T T N H T L I V H T T T

GACCCCCCAGGGCGCGCGCTCACCGGCCTCACCCGCACCGCACAAAACGA 4000  
D P P G G A A V T G L T R T A Q N E  
ACACCCCGGGCGCATCCACCTCATCGAAACCCACCCACCCACACCCAC 4050  
H P G R I H L I E T H E P H T F  
5 TCCCTCTACCCCACTCACCACCTCCACCAACCCACCTACGCTCACC 4100  
L P L T Q L T T L H Q P H L R L T  
AACAACACCCCTCCACACCCCGCACCTCACCCCATCACCACCCACCAAA 4150  
N N T L H T P H L T P I T T H H N  
CACCACCAACCCACCCCAACACCCACCCCTCAACCCCAACCCACGCCA 4200  
10 T T T T T P N T P P L N P N H A  
TCCTCATCACCGCGCGCTCCGGCACCCCTCGCGGCATCCTCGCTCGCC 4250  
I L I T G G S G T L A G I L A R H  
CTCAACACCCCGCACACCTACCTCCTCTCCCGCACACCACCCACCCAC 4300  
L N H P H T Y L L S R T P P P P T  
15 CACACCCCGGCACCCACATCCCTGCGACCTCACCGACCCACCCAAATCA 4350  
T P G T H I P C D L T D P T Q I  
CCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCTTCCACACC 4400  
T Q A L T H I P Q P L T G I F H T  
GCCGCCACCCCTCGACGACGCCACCCCTCACCAACCTCACCCCGCACACCT 4450  
20 A A T L D D A T L T N L T P Q H L  
CACCACACCCCTCCAACCCAAAGCGGACGCCGCTGGCACCTCCACCACC 4500  
T T T L Q P K A D A A W H L H H  
ACACCCAAAACCAACCCCTCACCCACTTCGTCTCTACTCCAGCGCCGCC 4550  
H T Q N Q P L T H F V L Y S S A A  
25 GCCACCCCTCGGCGAGCCCGGCCAAGCCAACTACGCCCGCCGCAACGCCTT 4600  
A T L G S P G Q A N Y A A A N A F  
CCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACCA 4600  
L D A L A T H R H T Q G Q P A T  
CCATCGCCTGGGGCATGTGGCACACCACCACACTCACCAGCCAACTC 4700  
30 T I A W G M W H T T T T L T S Q L  
ACCGACAGCGACCGGACCGCATCCGCCGCGCGGCTTCTGCGGATCTC 4750  
T D S D R I R R G G F L P I S  
GGACGACGAGGGCATGC  
D D E G M  
35

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
M R L Y E A A R R T G S P V V V  
40 CGGCGCGGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100  
A A A L D A P D V P L L R G L R  
GCGTACGACCGTCCGCGGTGCCGCGCTCCGGGAACGCTCTCTCGCCGACC 150  
R T T V R R A A V R E R S L A D  
GCTCGCGGTGCTGCCCCGACGACGAGCGCGCGGACGCTCCCTCGCGTTCC 200  
45 R S P C C P T T S A P T P P S R S  
TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250  
S W N S T A T V L G H L G A E D I  
CCCGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300  
P A T T T F K E L G I D S L T A  
50 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350  
V Q L R N A L T T A T G V R L N A  
ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400  
T A V F D F P T P R A L A A R L G  
CGACGAGCTGGCCGTTACCCGCGCGCCCGTCCGCGCCCGGACCGCGGCCA 450  
55 D E L A G T R A P V A A R T A A  
CCCGCGCCCGCGCACGACGACCGCTGGCGATCGTGGGATGGCTTCCCT 500  
T A A A H D E P L A I V G M A C R  
CTGCCCGGCGGGGTCCGCTCGCCACAGGAGCTGTGGCGTCTCGTCCGCTC 550  
L P G G V A S P Q E L W R L V A S

... SUBSTITUTE SHEET (RULE 26)

CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600  
G T D A I T E F P A D R G W D V  
ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650  
D A L Y D P D F D A I G K T F V R  
5 CACGGCGGCTTCCTCGACGGTSCGACCGGCTTCGACCGCGGCTTCCTCGG 700  
H G G F L D G A T G F D A A F F G  
GATCAGCCCCGCGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750  
I S P R E A L A M D P Q Q R V L  
TGGAGACGTCTGGGAGGCGTTGAAAGCGCGGGCATCACCCCGGACGCG 800  
10 L E T S W E A F E S A G I T P D A  
GCGCGGGGACGACACCGGCGTGTTCATCGGCGGCTTCCTACGGGTGTA 850  
A R G S D T G V F I G A F S Y G Y  
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGCCA 900  
G T G A D T N G F G A T G S Q T  
15 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950  
S V L S G R L S Y F Y G L E G P S  
GTCACGGTGCACACCGCCTGCTCGTCGTCACCTGGTCGCCCTGCACAGGC 1000  
V T V D T A C S S S L V A L H Q A  
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050  
20 G Q S L R S G E C S L A L V G G  
TCACGGTGATGGCGTCGCCCCGCGATTCTCGTCTGAGTTCTCCCGGACGCG 1100  
V T V M A S P G G F V E F S R Q R  
GGGCTCGCGCCGGACGGGCGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150  
G L A R D G R A K A F G A G A D G  
25 TACGAGCTTCGCGGAGGGCGCCGGTGCCTGGTGGTCGAGCGGCTCTCCG 1200  
T S F A E G A G A L V V E R L S  
ACGCGGAGCGCCACGGCCACACCGTCTCTGCCCTCGTACGCGGCTCCGCG 1250  
D A E R H G H T V L A L V R G S A  
GCTAACTCCGACGGCGCGTGAACGGTCTGTGCGCGCCGAACGGCCCTC 1300  
30 A N S D G A S N G L S A P N G P S  
CCAGGAACGCGTCATCCACAGGCCCTCGCGAACGCGAACTCACCCCCG 1350  
Q E R V I H Q A L A N A K L T P  
CCGATGTGACGCGGTGAGGGCGACGGCACCGGCACCCGCTCGGCGAC 1400  
A D V D A V E A H G T G T R L G D  
35 CCCATCGAGGCGCAGGCGTGTCTCGCGACGTACGGACAGGACCGGGCGAC 1450  
P I E A Q A L L A T Y G Q D R A T  
GCCCCGTGCTCGGCTCGCTGAAGTGAACATCGGGCACGCCAGGCCG 1500  
P L L L G S L K S N I G H A Q A  
CGTCAGGGGTGCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550  
40 A S G V A G I I K M V Q A I R H G  
GAACTGCCGCGACACTGCACGCGGACGCGTCCCGCACGTGACTG 1600  
E L P P T L H A D E P S P H V D W  
GACGGCCGGTGGCGTCTGAGCTCTGACGTGCGCCCGGCGGTGGCCGGGA 1650  
T A G A V E L L T S A R P W P G  
45 CCGGTGCGCCGCGCGCGCTGCCGTCTCGTCTGCGGTGAGCGGCACG 1700  
T G R P R R A A V S S F G V S G T  
AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTGCA 1750  
N A H I I L E A G P V K T G P V E  
GGCAGGAGCGATCGAGGCAGGACCGGTGAAGTAGGACCGGTGAGGCTG 1800  
50 A G A I E A G P V E V G P V E A  
GACCGCTCCCGCGGCGCGCGTCCGTCAGCACCGGGCGAAGACCTTCGCTG 1850  
G P L P A A P P S A P G E D L P L  
CTCGTGTGCGCGGCTTCGCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
L V S A R S P E A L D E Q I G R L  
55 GCGCGCTATCTCGACACCGGCGCGGCGTACCGGGCGGCGGTGGCGC 1950  
R A Y L D T G P G V D R A A V A  
AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCGTACTGCTCGGG 2000  
Q T L A R R T H F T H R A V L L G  
GACACCGTATCGGCGCTCCCCCGCGGACAGGCGGACGAAGTCTCTT 2050  
60 D T V I G A P P A D Q A D E L V F

SUBSTITUTE SHEET (RULE 26)

CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100  
V Y S G Q G T Q H P A M G E Q L  
CCGCCCGGTTCCCCGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTG 2150  
A A A F P V F A R I H Q Q V W D L  
5 CTGGATGTGCCCGATCTGAGGTGAACGAGACCGGTTACGCCCAGCCGGC 2200  
L D V P D L E V N E T G Y A Q P A  
CCTGTTCCGAATGCAGSTCGCTCTGTTCGGGCTGCTGGAATCGTGGGGTG 2250  
L F A M Q V A L F G L L E S W G  
10 TACGACCGGACCGGTTGATCGGCCATTCGGTGGGTGAGCTTGCGGCTGCG 2300  
T R P D A V I G H S V G E L A A A  
TATGTGTCCGGGGTGTGGTCTGGAGGATGCCTGCACTTTGGTGTCCGGC 2350  
Y V S G V W S L E D A C T L V S A  
GCGGGCTCGTCTGATGCAGGCTCTGCCCCGCGGGTGGGGTGATGGTCTGCTG 2400  
R A R L M Q A L P A G G V M V A  
15 TCCCGGTCTCGGAGGATGAGGCCCGGGCGGTGCTGGGTGAGGGTGTGGAG 2450  
V P V S E D E A R A V L G E G V E  
ATCGCCGCGGTCAACGGCCCGTCTGCTGGTGTCTCTCCGGTGTGAGGC 2500  
I A A V N G P S S V V L S G D E A  
CGCCGTGCTGACGGCCCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGA 2550  
20 A V L Q A A E G L G K W T R L A  
CCAGCCACGCGTTCCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC 2600  
T S H A F H S A R M E P M L E E F  
CGSGCGGTGCGCCGAAGGCTGACCTACCGGACGCGCAGGTCTCCATGGC 2650  
R A V A E G L T Y R T P Q V S M A  
25 CGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGG 2700  
V G D Q V T T A E Y W V R Q V R  
ACACGGTCCGGTTCCGGCGAGCAGGTGGCCTCGTACGAGGACGCGGTGTTT 2750  
D T V R F G E Q V A S Y E D A V F  
30 GTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTGCG 2800  
V E L G A D R S L A R L V D G V A  
GATGCTGCACGGCGACCAAGAAATCCAGGCCGCGATCGGCGCCCTGGCCC 2850  
M L H G D H E I Q A A I G A L A  
ACCTGTATGTCAACGGCGTCACGGTTCGACTGGCCCGCTCCTGGGCGAT 2900  
H L Y V N G V T V D W P A L L G D  
35 CTTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCA 2950  
A P A T R V L D L P T Y A F Q H Q  
GCGCTACTGGCTCGAGTCCGGTCCCCCGGCCACGGCCGACTCGGGCCACC 3000  
R Y W L E S A P P A T A D S G H  
CGTCTCTCGGCACCGGATCGCCGTGCGCGGGTTCGCGGGCCGGGTGTTT 3050  
40 F V L G T G V A V A G S P G R V F  
ACGGGTCCCGTGGCCCGCGGTGCGGACCGCGCGGTGTTTCATCGCCGAAT 3100  
T G P V P A G A D R A V F I A E L  
GGCGCTCGCCCGCCGCGACGCCACCGACTGCGCCACGGTCAACAGCTCG 3150  
A L A A A D A T D C A T V E Q L  
45 ACGTCACCTCCGTGCCCCGCGGATCCGCCCGCGCAGGGCCACCGCGCAG 3200  
D V T S V P G G S A R G R A T A Q  
ACCTGGGTGCGATGAACCCCGCCGCGACGGGCGGCGCGCTTACCGTCCA 3250  
T W V D E P A A D G R R R F T V H  
CACCCCGGTGCGGACGCCCCGTGGACGCTGCACGCCGAGGGGTTCTCC 3300  
50 T R V G D A P W T L H A E G V L  
GCCCCGGCCCGGTGCCCCAGCCCCGAAGCCGTGACACCGCCTGGCCCCCG 3350  
R P G R V P Q P E A V D T A W P P  
CCGGGCGCGGTGCCCCGCGACGGGCTGCCCGGGCGTGGCGACGCGCGGA 3400  
F G A V P A D G L P G A W R R A D  
55 CCAGGTCTTCGTGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCAC 3450  
Q V F V E A E V D S P D G F V A  
ACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGGACGGGAGCCGC 3500  
H P D L L D A V F S A V G D G S R  
CAGCCGACCGGATGGCGCGACCTCGCGGTGCACGCGTCCGACGCCACCGT 3550  
60 Q P T G W R D L A V H A S D A T V

GCTGCGCGCCTGCCTCACCCGCGCGGACAGTGGTGTGCTGGAGCTCGCCG 3600  
 L R A C L T R R D S G V V E L A  
 CCTTCGACGGTGCCGGAATGCCGGTGCTCACCCGCGGAGTCGGTGACGCTG 3650  
 A F C G A G M P V L T A E S Y T L  
 5 GCGGAGGTGCGGTGCGGACGCGGATCCGACGAGTCGGACGGTCTGCTTCG 3700  
 G E V A S A G G S D E S D G L L R  
 GCTTGAGTGGTTGCCGGTGCGGAGGCCACTACGACGGTGCCGACGAGC 3750  
 L E W L P V A E A H Y D G A D E  
 TGCCCGAGGGGTACACCCTCATCACCGCCACACACCCCGACGACCCCGAC 3800  
 10 L P E G Y T L I T A T H F D C P D  
 GACCCCAACCAACCCCAACACACCCACACGACCCCAACACAAACAC 3850  
 D P T N P H N T P T R T H T Q T T  
 ACGCGTCTTCACCGCCCTCCAACACCACCTCATCACCAACCAACACACCC 3900  
 R V L T A L Q H H L I T T N H T  
 15 TCATCGTCCACACCACCGACCCCGAGGCGCGCGCTCACCGGCCTC 3950  
 L I V H T T T D P P G A A V T G L  
 ACCCGCACCGGCACAAACGAACACCCCGGCGCGCATCCACCTCATCGAAAC 4000  
 T R T A Q N E H P G R I H L I E T  
 CCACACCCCAACCCCACTCCCTCACCCTCACCACCCCTCCACC 4050  
 20 H H P H T P L P L T Q L T T L H  
 AACCCCACTACGCTCACCAACAACACCTCCACACCCCACTCACC 4100  
 Q P H L R L T N N T L H T P H L T  
 CCCATCACCAACCACCAACACCACCAACACCCCAACACCCCAAC 4150  
 P I T T H H N T T T T T P N T P P  
 25 CCTCAACCCCAACACGCGCATCCTCATCACCGGCGGCTCCGGCACCCCTCG 4200  
 L N P N H A I L I T G G S G T L  
 CCGGCATCTCGCCCGCACCTCAACCAACCCCAACCTACCTCCTCTCC 4250  
 A G I L A R H L N H P H T Y L L S  
 CGCACACCAACCAACCCCAACACCCCGGACCCACATCCCTGCGACCT 4300  
 30 R T P P P P T T P G T H I P C D L  
 CACCGACCCCAACCAATCACCAAGCCCTCACCAACATACCAACCAACCC 4350  
 T D P T Q I T Q A L T H I P Q P  
 TCACCGGATCTTCACACCGCGCGCACCCCTCGACGAGCCACCTCACC 4400  
 L T G I F H T A A T L D D A T L T  
 35 AACCTCACCCCAACACCTCACCAACACCCCTCAACCCCAAGCGGACGC 4450  
 N L T P Q H L T T T L Q P K A D A  
 CGCCTGGCACCTCCACCAACCAACCAACCCCTCACCACTTCG 4500  
 A W H L H H H T Q N Q P L T H F  
 TCCTCTACTCCAGCGCGCGCCACCCCTCGGCAGCCCGGCAAGCCAAC 4550  
 40 V L Y S S A A T L G S P G Q A N  
 TACGCGCGCGCAACGCCTTCCTCGACGCGCTCGCCACCAACCGCCACAC 4600  
 Y A A A N A F L D A L A T H R H T  
 CCAAGGACAACCCGCCACCAACATCGCCTGGGGCATGTGGCACACCA 4650  
 Q G Q P A T T I A W G M W H T T  
 45 CCACACTCACCACTCAACCGACAGCGACCGCGACCGCATCCGCGCG 4700  
 T T L T S Q L T D S D R D R I R R  
 GCGCGCTTCCTGCGCATCTCGGACGACGAGGCGATGC  
 G G F L P I S D D E G M

50 The *NheI*-*XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module  
 13 of rapamycin is shown below.

CCATGCGSCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50  
 M R L Y E A A R R T G S P V V V  
 GCGGCCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGG33CTGCG 100  
 55 A A A L D D A P D V P L L R G L R  
 GGTACGACCGTCCGCGCTGCGCGCGTCCGGAACGCTCTCTCCCGGACC 150  
 R T T V R R A A V R E R S L A D  
 GCTCGCCGTGCTGCCGACGACGAGCGCGCGGACGCTCCCTCGCGTTTCG 200  
 R S F C C P T T S A P T P P S R S

	TCCTGGAACAGCACCCGCCACCGTGTCTCGGCCACCTGGGCGCCGAAGACAT	250
	S W N S T A T V L G H L G A E D I	
	CCCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG	300
	F A T T T F K E L G I D S L T A	
5	TCCAGCTGCGCAACCGCGCTGACCAACGGCGACCGGCGTACGCCTCAACGCC	350
	V Q L R N A A C L T T A T G V R L N A	
	ACAGCGGTCTTCGACTTTCCGACCGCGCGCTCGCCGCGAGACTCGG	400
	T A V F D F P T P R A L A A R L G	
10	CGACGAGCTGGCCCGTACCCGCGCGCCCGTCTCGCGGCCCGGACCGCGGCCA	450
	D E L A G T R A P V A A R T A A	
	CCCGCGCCGCGCACGACGAACCGCTGGCGATCGTGGGCGATGGCCTGCCGT	500
	T A A A H D E P L A I V G M A C R	
	CTGCCGGGGGGGTCCGCTCGCCACAGGAGCTGTGGCGTCTCGTCCGCTC	550
	L P G G V A S P Q E L W R L V A S	
15	CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG	600
	G T D A I T E F P A D R G W D V	
	ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG	650
	D A L Y D P D P D A I G K T F V R	
20	CACGGCGGTTCTCTGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG	700
	H G G F L D G A T G G F D A A F F G	
	GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGACGAACGGGTGCTCC	750
	I S P R E A L A M D P Q Q R V L	
	TGGAGACGTCTTGGGAGGCGTTTCGAAAGCGCGGGCATCACCCCGGACGCG	800
25	L E T S W E A F E S A G I T P D A	
	GCGCGGGGCAAGCGACACCGCGGTGTTCTCGGCGGTTCTCTACGGGTA	850
	A R G S D T G V F I G A F S Y G Y	
	CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGACACCA	900
	G T G A D T N G F G A T G S Q T	
30	GCGTGTCTTCGGCGCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG	950
	S V L S G R L S Y F Y G L E G P S	
	GTCACGGTTCGACACCGCTGCTCGTCTGCTACTGGTTCGCCCTGCACCAGGC	1000
	V T V D T T A C S S S L V A L H Q A	
	AGGGCAGTCCCTGCGTCCGGCGAATGCTCGCTCGCCCTGGTGGCGGGT	1050
	G Q S L R S G E C S L A L V G G	
35	TCACGGTGATGGCGTTCGCCCGCGGATTCTGTCGAGTTCTCCCGGCAGCGC	1100
	V T V M A S P G G F V E F S R Q R	
	GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG	1150
	G L A P D G R A K A F G A G A D G	
40	TACGAGCTTCGCGAGGGCGCGGTGCCCTGGTGGTCGAGCGGCTCTCCG	1200
	T S F A E G A G A L V V E R L S	
	ACGCGGAGCGCCACGGCCACACCGTCTCTGCCCTCGTACGCGGCTCCGCG	1250
	D A E R H G H T V L A L V R G S A	
	GCTAACTCCGACGGCGCGTGAACGGTCTGTCTGGCGCCGAACGGCCCCCTC	1300
	A N S D G A S N G L S A P N G P S	
45	CCAGGAACCGGTATCCACAGGCCCTCGCGAACGCGAACTCACCCCCG	1350
	Q E R V I H Q A L A N A K L T T	
	CCGATGTGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC	1400
	A D V D A V E A H G T G T R L G D	
50	CCCATCGAGGGCGAGGCCCTGCTCGCGACGTACGGACAGGACCGGGCGAC	1450
	F I E A Q A L L A T Y G Q D R A T	
	CCCCCTGCTGCTCGGCTCGGTGAAGTCGAACATCGGGCACGCCACGGCCG	1500
	F L L G S L K S N I G H A Q A	
	CGTCAGGGGTTCGGCGGATCAAGATGGGTGACAGGCCATCCGGCACGGG	1550
	A S G V A G I I K M V Q A I R H G	
55	UAAGTCCCGCGACACTGCACGCGGACGAGCCGTGCGCGCACGTCGACTG	1600
	F L P P T L H A D E P S P H V D W	
	TACCGCGGCTCGCGTTCGAGCTCTGACGTTCGGCCCGGCGGTGGCCGGGGA	1650
	T A G A V E L L T S A R P W F G	
60	CCGGTTCGCCCGCGCGCTGCCGTCTCGTCTGCTCGGCGTGGAGCGGCACG	1700
	T G R P R R A A V S S F G V S G T	

**SUBSTITUTE SHEET (RULE 26)**

AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750  
N A H I I L E A G P V K T G P V E  
GGCAGGAGCGATCGAGGCAGGACCGGTCTGAAGTAGGACCGGTCTGAGGCTG 1800  
A G A I E A G P V E V G P V E A  
5 GACCGCTCCCCGCGCGCCGCGCTCAGCACCGGGCGAAGACCTTCCGCTG 1850  
G P L P A A P P S A P G E D L P L  
CTCGTGTGCGGCGGTTCCTCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900  
L V S A R S P E A L D E Q I G R L  
10 GCGCGCCTATCTCGACACCGGCGCGGGCGTCTGACCGGGCGGCGCTGGCGC 1950  
R A Y L D T G P G V D R A A V A  
AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCCCTACTGCTCGGG 2000  
Q T L A R R T H F T H R A V L L G  
GACACCGTCTCGGCGCTCCCCCGCGGACCGAGGCGGACGAACCTCGTCTT 2050  
D T V I G A P P A D Q A D E L V F  
15 CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100  
V Y S G Q G T Q H P A M G E Q L  
CCGATTCTGTCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCG 2150  
A D S S V V F A E R M A E C A A A  
TTGCGCGAGTTCGTGGACTGGGATCTGTTACGGTTCCTGGATGATCCGGC 2200  
20 L R E F V D W D L F T V L D D P A  
GGTGGTGGACCGGTTGATGTGGTCCAGCCGCTTCTGGGCGATGATGG 2250  
V V D R V D V V Q P A S W A M M  
TTTCCCTGGCCGCGGTGTGGCAGGCGGCGGTGTGCGGCGGATGCGGTG 2300  
V S L A A V W Q A A G V R P D A V  
25 ATCGGCCATTTCGACGGGTGAGATCGCCGAGCTTGTGTGGCGGGTGCAGT 2350  
I G H S Q G E I A A A C V A G A V  
GTCACTACGCGATGCCGCGCGGATCGTGACCTTGCGCAGCCAGGCGATCG 2400  
S L R D A A R I V T L R S Q A I  
CCCGGGGCGCTGGCGGGCGGGGCGCGATGGCATCCGTGCGCCCTGCCCGCG 2450  
30 A R G L A G R G A M A S V A L P A  
CAGGATGTGAGCTGGTTCGACGGGGCGCTGGATCGCCGCCACAACGGGGC 2500  
Q D V E L V D G A W I A A H N G P  
CGCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTCTGACCATGTCTCA 2550  
A S T V I A G T P E A V D H V L  
35 CCGCTCATGAGGCACAAGGGGTGCGGGTGCAGCGGATCACCGTCTGACTAT 2600  
T A H E A Q G V R V R R I T V D Y  
GCCTCGCACACCCCGCACCTCGAGCTGATCCGCGACGAACCTACTCGACAT 2650  
A S H T P H V E L I R D E L L D I  
CACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTGACCG 2700  
40 T S D S S S Q T P L V P W L S T  
TGGACGGCACCTGGGTCTGACAGCCCGCTGGACGGGGAGTACTGGTACCGG 2750  
V D G T T W V D S P L D G E Y W Y R  
AACCTGCGTGAACCGGTCTGGTTTCCACCCCGCGTCTAGCCAGTTGCAGGC 2800  
N L R E P V G F H P A V S Q L Q A  
45 CCAGGGGACACCGTGTTCGTCTGAGGTCTAGCGCCAGCCCGGTGTGTGTC 2850  
Q G D T V F V E V S A S P V L L  
AGGCGATGGACGACGATGTCTGTCACGGTTGCCACGCTGCGTCTGTGACGAC 2900  
Q A M D D D V T V A T L R R D D  
GGCGACGCCACCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGG 2950  
50 S D A T R M L T A L A Q A Y V H G  
CCTCACCGTCTGACTGGCCCGCCATCCTCGGCACCAACACAACCCGGGTAC 3000  
V T V D W P A I L G T T T T R V  
TGGACCTTCCGACCTACGCCCTTCCACACAGCGGTACTGGCTCGAGTCG 3050  
L D L P T Y A F Q H Q R Y W L E S  
55 GCTCCCCCGGCGGCGGACTCGGGCCACCCGTCCTCGGCACCGGAGT 3100  
A P P A T A D S G H P V L G T G V  
CGCCGTCGCGGGGTGCGCGGGCGGGGTGTTACGGGTCCCGTCCCCGCCG 3150  
A V A G S P G R V F T G P V P A  
GTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCCGCCGAC 3200  
60 G A D R A V F I A E L A L A A A D



3250  
A T D C A T V E Q L D V T S V P G  
CGGATCCGCCCCGCGCAGGGCCACCGCGCAGACCTGGGTGCGATGAACCCG 3300  
5 S A R E R A T A Q T W V D E F  
CCGCCGACGGGCGGCGCGCTTCACCGTCCACACCCGCGTCCGGCGACGCC 3350  
A A D G R R R F T V H T R V G D A  
CCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCGCGCGCGTGGCCCA 3400  
P W T L H A E G V L R P G R V P Q  
10 GCCCGAAGCCGTCGACACCGCCTGGCCCCCGCGGGCGCGGTGCCCGCGG 3450  
P E A V D T A W P P P G A V P A  
ACGGGCTGCCCCGGGGCGTGGCGACGCGCGGACCAGGTCTTCGTGGAAGCC 3500  
D G L P G A W R R A D Q V F V E A  
GAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGC 3550  
E V D S P D G F V A H P D L L D A  
15 GGTCTTCTCCGCGGTCCGGCGACGGGAGCCGCCAGCCGACCGGATGGCGCG 3600  
V F S A V G D G S R Q P T G W R  
ACCTCGCGGTGCACGCGTCCGGACGCCACCGTGTGCGCGCCTGCCTCACC 3650  
D L A V H A S D A T V L R A C L T  
CGCCGCGACAGTGGTGTCTGTGGAGCTCGCCGCTTCGACGGTGGCGGAAT 3700  
20 R R D S G V V E L A A F D G A G M  
GCCGGTGTCTACCGCGGAGTCCGGTGACGCTGGGCGAGGTCCGCTCGGCAG 3750  
P V L T A E S V T L G E V A S A  
GCGGATCCGACGAGTCCGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTG 3800  
G G S D E S D G L L R L E W L P V  
25 GCGGAGGCCCCACTACGACGGTCCCGACGAGCTGCCCGAGGGCTACACCT 3850  
A E A H Y D G A D E L P E G Y T L  
CATCACCGCCACACACCCCGACGACCCCGACGACCCCAACCCCAACA 3900  
I T A T H P D D P D D P T N P H  
ACACACCCACACGCACCCACACACAAACCACACGCGTCTCACCAGCCCTC 3950  
30 N T P T R T H T Q T T R V L T A L  
CAACACCCTCATCACCAACCAACCCCTCATCGTCCACACCACCAC 4000  
Q H H L I T T N H T L I V H T T T  
CGACCCCCCAGGCGCGCGCTCACCGGCCTCACCGCACCGCACAAAACG 4050  
D P P G A A V T G L T R T A Q N  
35 AACACCCCGCGCGCATCCACCTCATCGAAACCCACACCCCAACACCCCA 4100  
E H P G R I H L I E T H H P H T P  
CTCCCCCTCACCAACTCACACCCCTCCACCAACCCCACTACGCTCACC 4150  
L P L Q L T T L H Q P H L R L T  
CAACAACACCTCCACACCCCTCACCCCTCACCCCTCACCCACCAACA 4200  
40 N N T L H T P H L T P I T T H H  
ACACCACCAACCAACCCCAACACCCCAACCCCTCAACCCCAACCAACG 4250  
N T T T T T P N T P P L N P N H A  
ATCCTCATCACCGGCGGCTCCGGCACCCCTCGCGGCATCCTCGCCCGCCA 4300  
I L I T G G S G T L A G I L A R H  
45 CCTCAACACCCCAACACCTACCTCCTCTCCGCGACACCAACCAACCCCA 4350  
L N H P H T Y L L S R T P P P P  
CCACACCCCGGCACCCACATCCCTGCGACCTCACCGACCCCAACCAATC 4400  
T T P G T H I P C D L T D P T Q I  
ACCCAAGCCCTCACCAACATACCAACCCCTCACCGGCATCTTCCACAC 4450  
50 T Q A L T H I P Q P L T G I F H T  
CGCGCGCACCCCTCGACGACGCCACCCCTCACCAACCTCACCCCAACACC 4500  
A A T L D D A T L T N L T P Q H  
TCACCAACCCCTCCAAACCCAAAGCCGACGCGCGCTGGCACCTCCACCAC 4550  
L T T T L Q P K A D A A W H L E H  
55 CACACCCCAACCAACCCCTCACCACTTCGTCTCTACTCCAGCGCGCGC 4600  
H T Q N Q P L T H F V L Y S S A A  
CCCCACCCCTCGGCAACCCCGGCAAGCCAACTACGCGCGCGCAACCCCT 4650  
A T L G S P G Q A N Y A A A N A  
TCCTCGACGCCCTCGCCACCCACCGCCACACCAAGGACAACCCGCCACC 4700  
60 F L D A L A T H R H T Q G Q P A T

SUBSTITUTE SHEET (RULE 26)

ACCATCGCCTGGGGCATCTGGCACCACCACCACACTCACCAGCCAACT 4750  
 T I A W G M W H T T T T L T S Q L  
 CACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCCTGCCGATCT 4800  
 T C S C R C R I R R C G F L P I  
 5 CCGACGACGAGGGCATGC  
 S D D E G M

### Example 3

#### Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

10 The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from *Streptomyces* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding  
 15 sequences have been replaced by either the *rapAT3* (the AT domain from module 3 of the rapamycin PKS), *rapAT12*, *eryAT1* (the AT domain from module 1 of the erythromycin (DEBS) PKS), or *eryAT2* coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the *rapAT12* replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other  
 20 derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites *SacI* and *SphI* (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising  
 25 module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique *SacI* and *SphI* restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique *Bgl* II and *Nsi* I sites by ligation to synthetic linkers (described in the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8  
 30 sequences were then amplified using primers, described above, that introduced either an *Avr* II site or an *Nhe* I site at two different KS/AT boundaries and an *Xho* I site at the AT/DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated  
 35 to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the *Bam* HI and *Pst* I sites of the

KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Heterologous AT	Enzyme	Location of Engineered Site
FK-506 AT8 (hydroxymalonyl)	<i>AvrII</i>	GGCCGT <u>ccgcgc</u> CGTGCGGCGGTCTCGTCGTTTC
	<i>NheI</i>	G R P R R A A V S S F ACCCAGCATCCCGCGATGGGTGAGCG <u>gcgcgc</u> C
	<i>XhoI</i>	T Q H P A M G E R L A TACGCCTTCCAGCGGCGGCCCTACTGG <u>gcgcgc</u> g
rapamycin AT3 (methylmalonyl)	<i>AvrII</i>	GACCGG <u>ccccgc</u> CGGGCGGGCGGTGTCGTCCTTC
	<i>NheI</i>	D R P R R A G V S S F TGGCAGTGGCTGGGGATGGGCAGTGC <u>cctgcgc</u> G
	<i>XhoI</i>	W Q W L G M G S A L R TACGCCTTCCAACACCAGCGGTACTGG <u>gcgcgc</u> g
rapamycin AT12 (malonyl)	<i>AvrII</i>	GGCCGA <u>gcgcgc</u> CGGGCAGGCGGTGTCGTCCTTC
	<i>NheI</i>	G R A R R A G V S S F TCGCAGCGTGCTGGCATGGGTGAGGA <u>actgcgc</u> C
	<i>XhoI</i>	S Q R A G M G E E L A TACGCCTTCCAGCACCAGCGCTACTGG <u>gcgcgc</u> g
DEBS AT1 (methylmalonyl)	<i>AvrII</i>	GCGCGA <u>ccccgc</u> CGGGCGGGGGTCTCGTCGTTTC
	<i>NheI</i>	A R P R R A G V S S F TGGCAGTGGGCGGGCATGGCCGTCGA <u>cctgcgc</u> C
	<i>XhoI</i>	W Q W A G M A V D L L TACCCGTTCCAGCGCGAGCGCGTCTGG <u>gcgcgc</u> g
DEBS AT2 (methylmalonyl)	<i>AvrII</i>	GACGGG <u>gcgcgc</u> CGGGCAGGTGTGTCGGCGTTTC
	<i>NheI</i>	D G V R R A G V S A F GCCCAGTGGGAAGGCATGGCGCGGGA <u>gttgcgc</u> G
	<i>XhoI</i>	A Q W E G M A R E L L TATCCTTTCCAGGGCAAGCGGTTCTGG <u>gcgcgc</u> g

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

CCGGCGCCGTCGAACTGCTGACGTCGGCCCCGGCGTGGCCCCGAGACCGACCGGccacggc  
 A G A V E L L T S A R P W P E T D R P R  
 GTGCGCGCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCCACGTCATCCTGGAGGCGG  
 P A A V S S F G V S G T N A H V T L E A  
 5 GACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGACCTTCCCTGCTGGTGTGCGG  
 G P V T E T P A A S P S G D L P L L V S  
 CACGCTCACCGGAAGCGCTCGACGAGCAGATCCGCGGACTGCGCGCCTACCTGGACACCA  
 A R S P E A L D E Q I R R L R A Y L D T  
 CCGCGGACGTCGACCGGGTGGCCGTGGCACAGACGCTGGCCCGGCGCACACACTTCGCCC  
 10 T P D V D R V A V A Q T L A R R T H F A  
 ACCGCGCGGTGCTGCTCGGTGACACCGTTCATCACACACCCCCGCGGACCGGCCCCGACG  
 H R A V L L G D T V I T T P P A D R P D  
 AACTCGTCTTCGTCTACTCCGGCCAGGGCAGCCAGCATCCCGCGATGGGCGAGCgctcg  
 E L V F V Y S G Q G T Q H P A M G E Q L  
 15 CCGCGCCCATCCCGTGTTCGCGGACGCCTGGCATGAAGCGCTCCCGCGCTTGACAACC  
 A A A H P V F A D A W H E A L R R L D N

The sequences shown below provide the location of the AT/DH boundary chosen in  
 the FK-520 module 8 coding sequences. The region where an *Xho*I site was engineered is  
 indicated by lower case and underlining.

TCCTCGGGGCTGGGTACGGCACGACCGCGGATGTGCCCGCGTACGCGTTCCAACGGCGGC  
 I L G A G S R H D A D V P A Y A F Q R R  
 ACTACTGGgacgagTCCGCACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGTGGGCT  
 25 H Y W I E S A R P A A S D A G H P V L G

The sequences shown below provide the location of the KS/AT boundaries chosen  
 in the FK-506 module 8 coding sequences. Regions where *Avr*II and *Nhe*I sites were  
 engineered are indicated by lower case and underlining.

TCGGCCAGGCGGTGGCCGCGGACCGGCCGTccgcccCGTGCGGCGGTCTCGTCTCGTTCGGG  
 30 S A R P W P R T G R P R R A A V S S F G  
 GTGAGCGGCACCAACGCCCACATCCTTGGAGGCGGACCCGACCAGGAGGAGCCGTG  
 V S G T N A H I I L E A G P D Q E E P S  
 GCAGAACCGGCCGGTGACCTCCCGCTGCTCGTGTGCGCACGGTCCCCGGAGGCACTGGAC  
 A E P A G D L P L L V S A R S P E A L D  
 35 GAGCAGATCGGGCGCCTGCGCGACTATCTCGACGCGCCCCCGGCGTGGACCTGGCGGCC  
 E Q I G R L R D Y L D A A P G V D L A A  
 GTGGCGCGGACACTGGCCACGCGTACGCACTTCTCCACCGCGCCGTACTGCTCGGTGAC  
 V A R T L A T R T H F S H R A V L L G D  
 ACCGTCTACACCGCTCCCCCGTGGAACAGCGGGCGAGCTCGTCTTCTGCTACTCGGGA  
 40 T V I T A P P V E Q P G E L V F V Y S G  
 CAGGGCACCCAGCATCCCGCGATGGGTGAGCGgctcgcCGCAGCCTTCCCGTGTTCGCC  
 Q G T Q H P A M G E R L A A A F P V F A  
 GACCCGGACGTACCCGCCTACGCCTTCCAGCGGCGGCCCTACTGGATCGAGTCCGCGCCG  
 45 D P D V P A Y A F Q R R P Y W I E S A P

The sequences shown below provide the location of the AT/DH boundary chosen in  
 the FK-506 module 8 coding sequences. The region where an *Xho*I site was engineered is  
 indicated by lower case and underlining.

GACCCGGACGTACCCGCCTACGCCTTCCAGCGGCGGCCCTACTGGgacgagTCCGCGCCG  
 50 D P D V P A Y A F Q R R P Y W I E S A P

#### Example 4

Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506 and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or methyl. These derivatives are produced in recombinant host cells of the invention that  
 5 express recombinant PKS enzymes the produce the derivatives. These recombinant PKS enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the present invention provides recombinant PKS enzymes in which the AT domains of both modules 7 and 8 have been changed. The table below summarizes the various compounds  
 10 provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
15	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound -- FK-506
	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
20	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
25	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound -- FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
30	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5

Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and/or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module.

Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domains but also those in which one of the modules is converted to an ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

#### Example 6

##### Neurotrophic Compounds

The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation.

These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of

FK-506. The 18-hydroxy compounds of the invention can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45  $\mu$ L of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64  $\mu$ L) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with brine, and the organic phase dried over magnesium sulfate. Removal of solvent *in vacuo* and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53  $\mu$ L of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-18-hydroxyl derivatives of the compounds described in Examples 1 - 4.

Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai *et al.*, Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, *FEBS Letters* 316(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[S]-OH and C18-[R]-OH enantiomers, with the R enantiomer showing a somewhat lower IC<sub>50</sub>, which may be preferred in some applications. See Kawai *et al.*, *supra*. Another preferred protocol is described in Umbreit and Sharpless, 1977, JACS 99(16): 1526-28, although it may be preferable to use 30 equivalents each of

SeO<sub>2</sub> and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

5 All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of illustration and not limitation of the following claims.



Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthetase, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.

5

2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

10

3. The isolated nucleic acid of claim 1 that encodes an open reading frame, said open reading frame comprising coding sequences for two or more extender modules, each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

15

4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

20

5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.

6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.

25

7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid pKOS065-M21.

30

8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase, FK-506 polyketide synthase, or erythromycin polyketide synthase.

9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.

10. The method of claim 9, wherein said host cell is a *Streptomyces* host cell.

11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.

12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.

13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

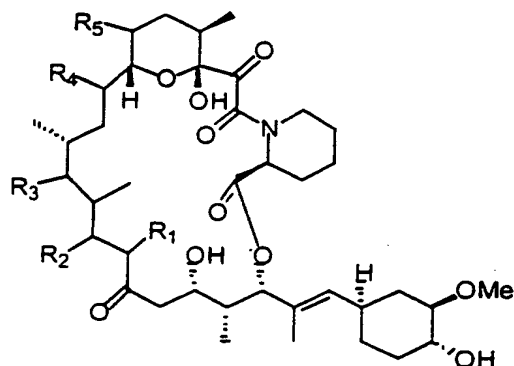
15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.

16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.

17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

18. A polyketide having the structure

5

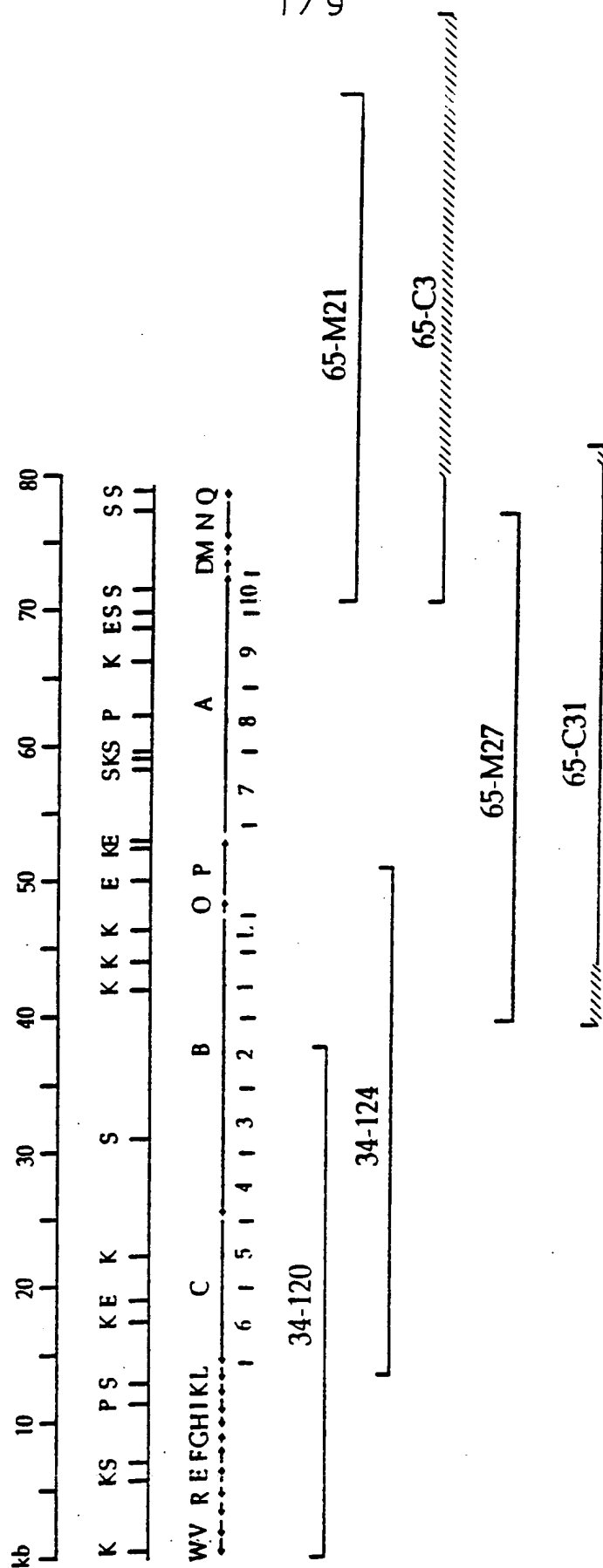


wherein, R<sub>1</sub> is hydrogen, methyl, ethyl, or allyl; R<sub>2</sub> is hydrogen or hydroxyl, provided that when R<sub>2</sub> is hydrogen, there is a double bond between C-20 and C-19; R<sub>3</sub> is hydrogen or hydroxyl; R<sub>4</sub> is methoxyl, hydrogen, methyl, or ethyl; and R<sub>5</sub> is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.

19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.

15

20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.



**FIG. 1**

2/9

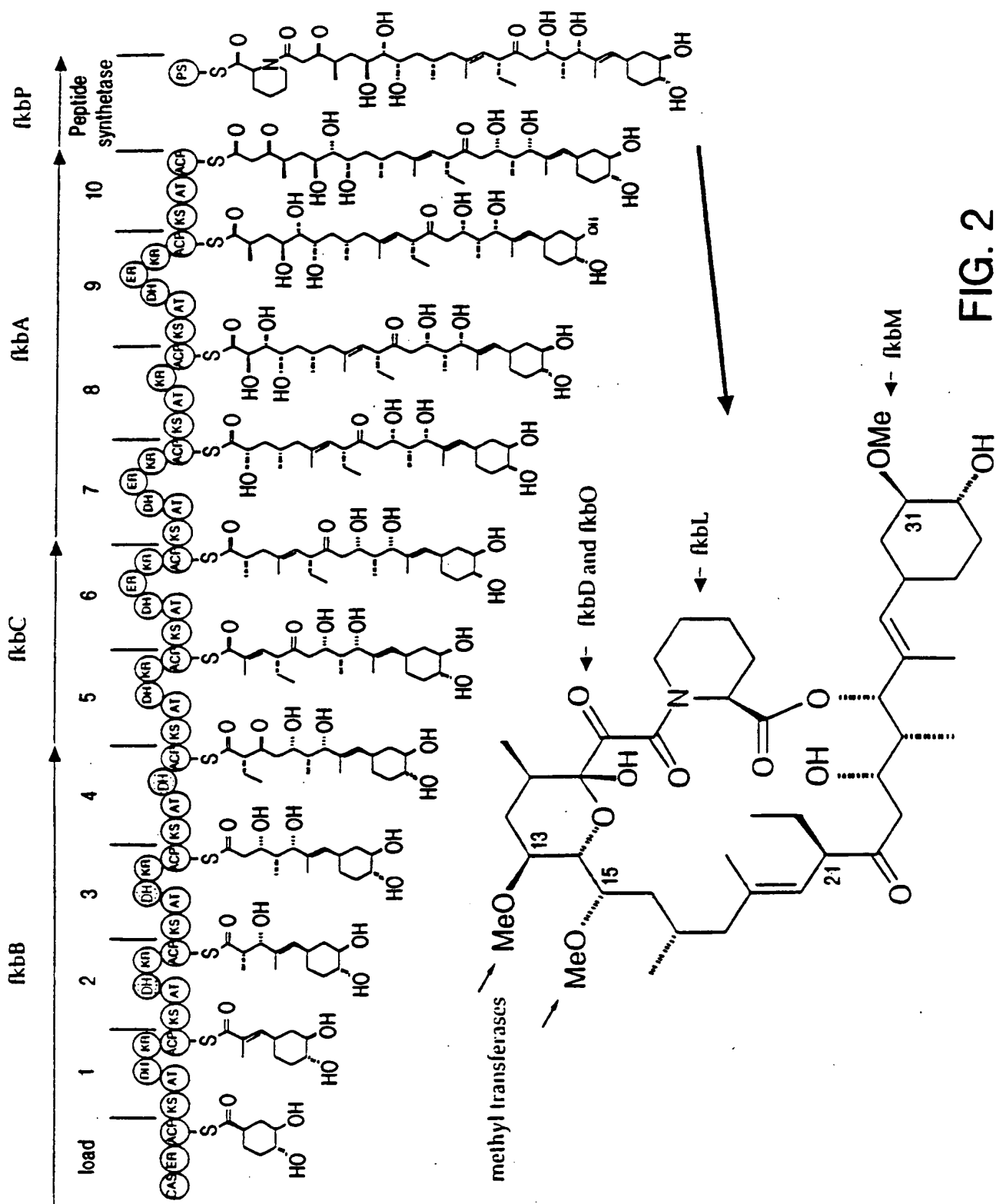


FIG. 2

SUBSTITUTE SHEET (RULE 26)

3/9

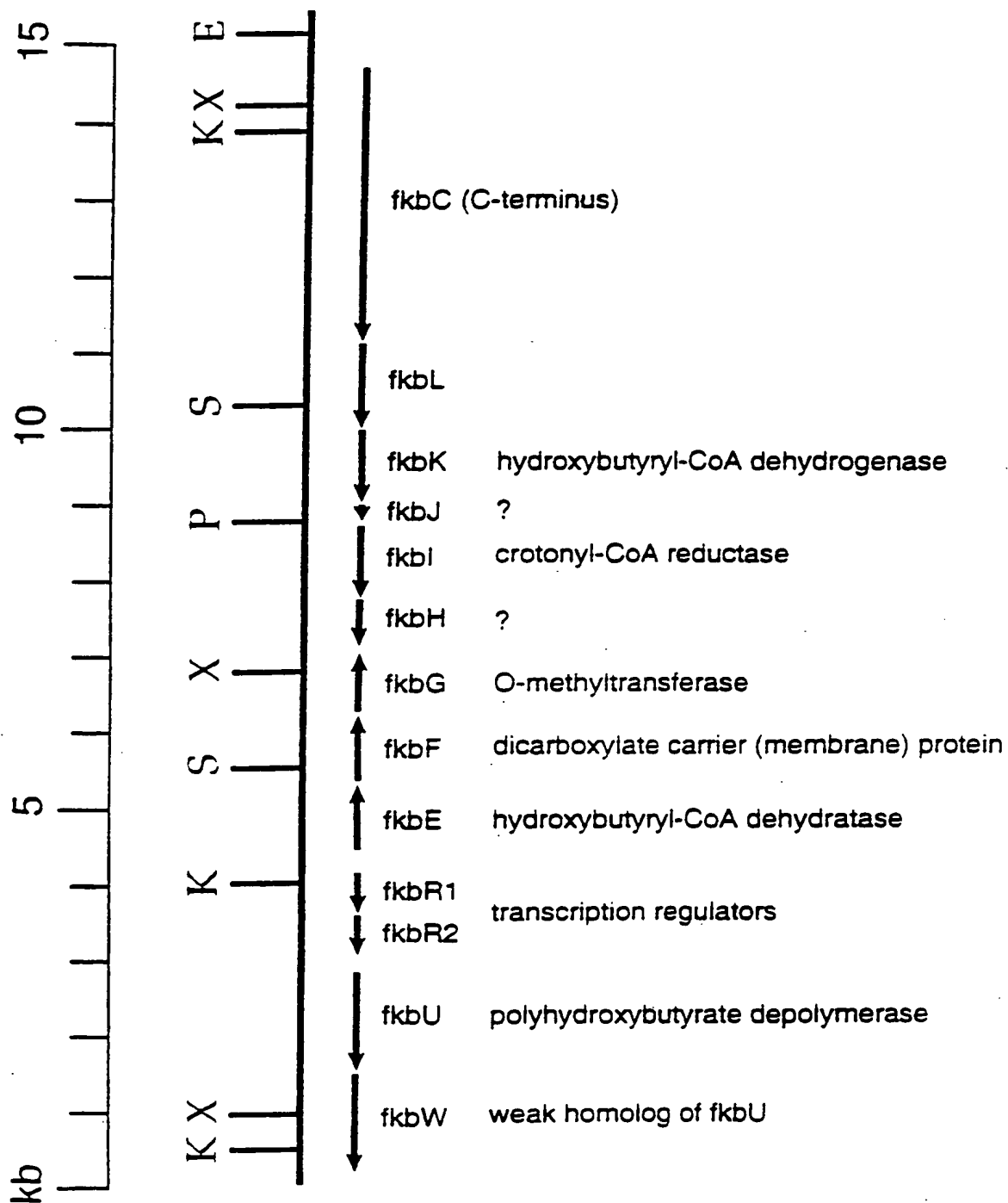


FIG. 3

SUBSTITUTE SHEET (RULE 26)

4 / 9

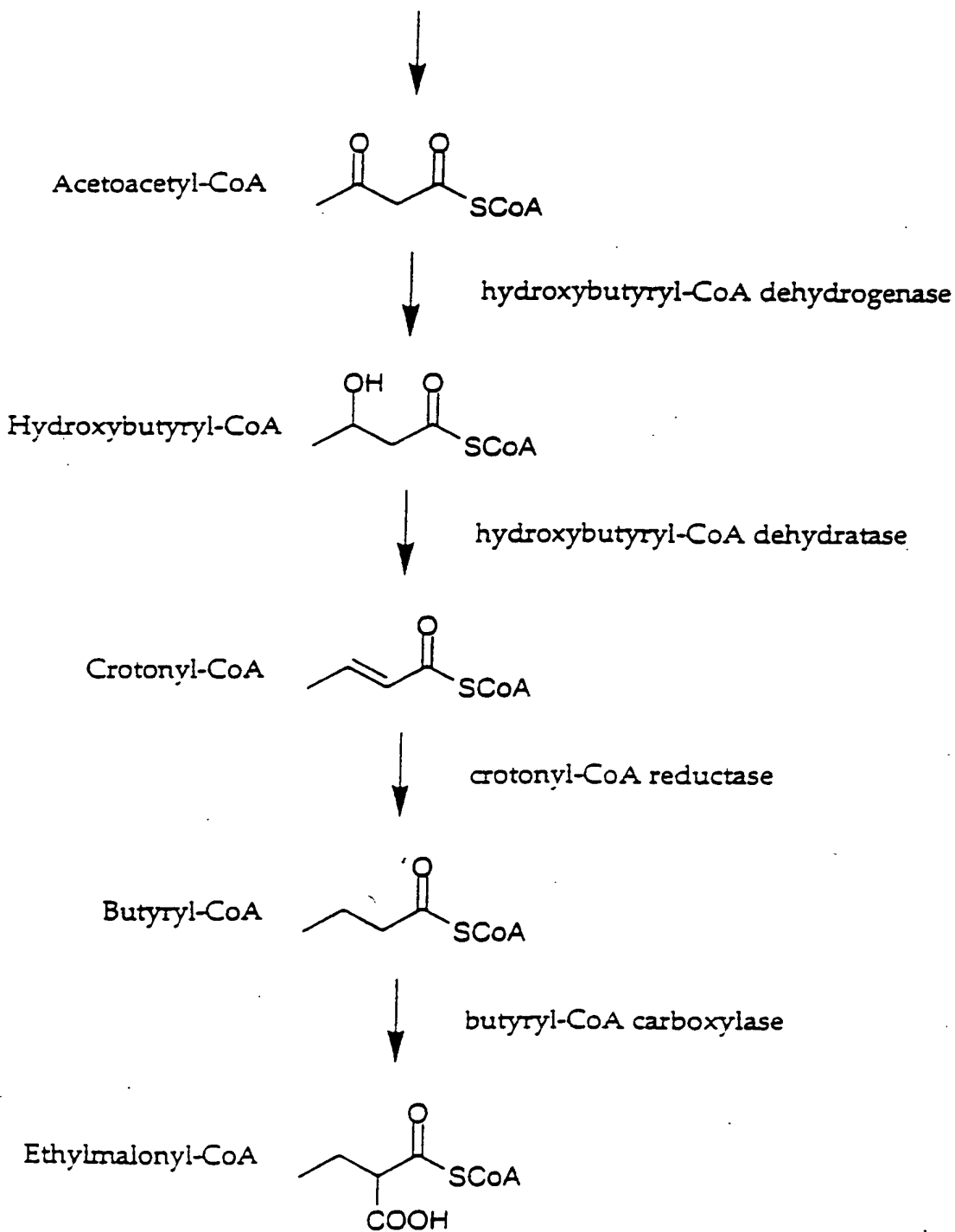
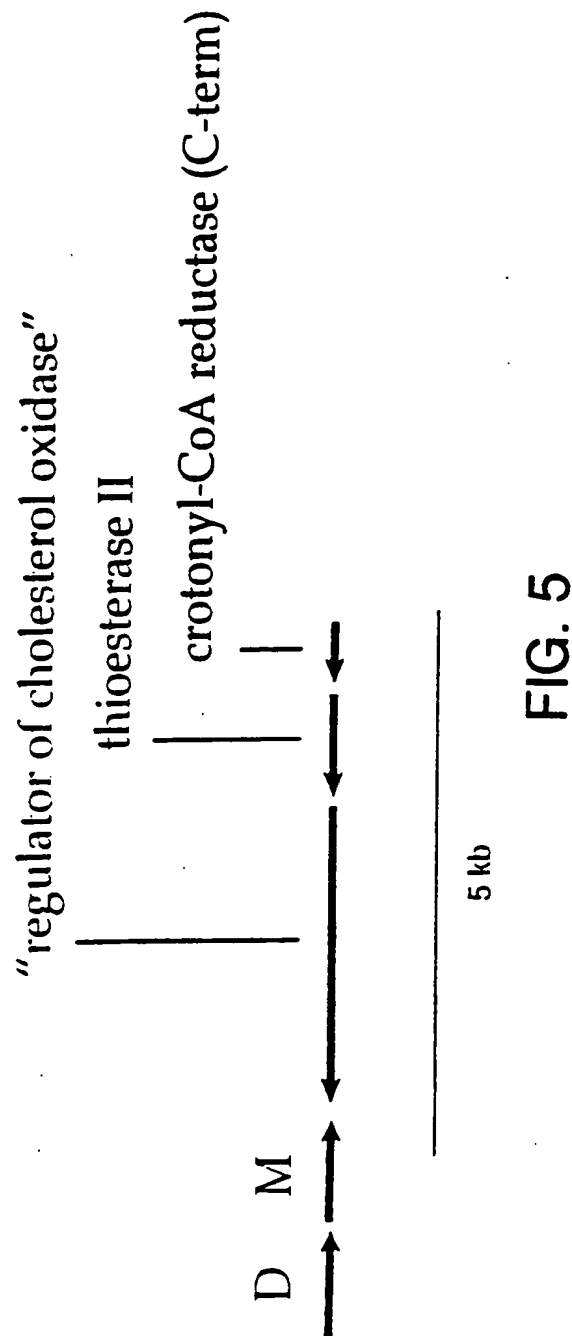


FIG. 4





6/9

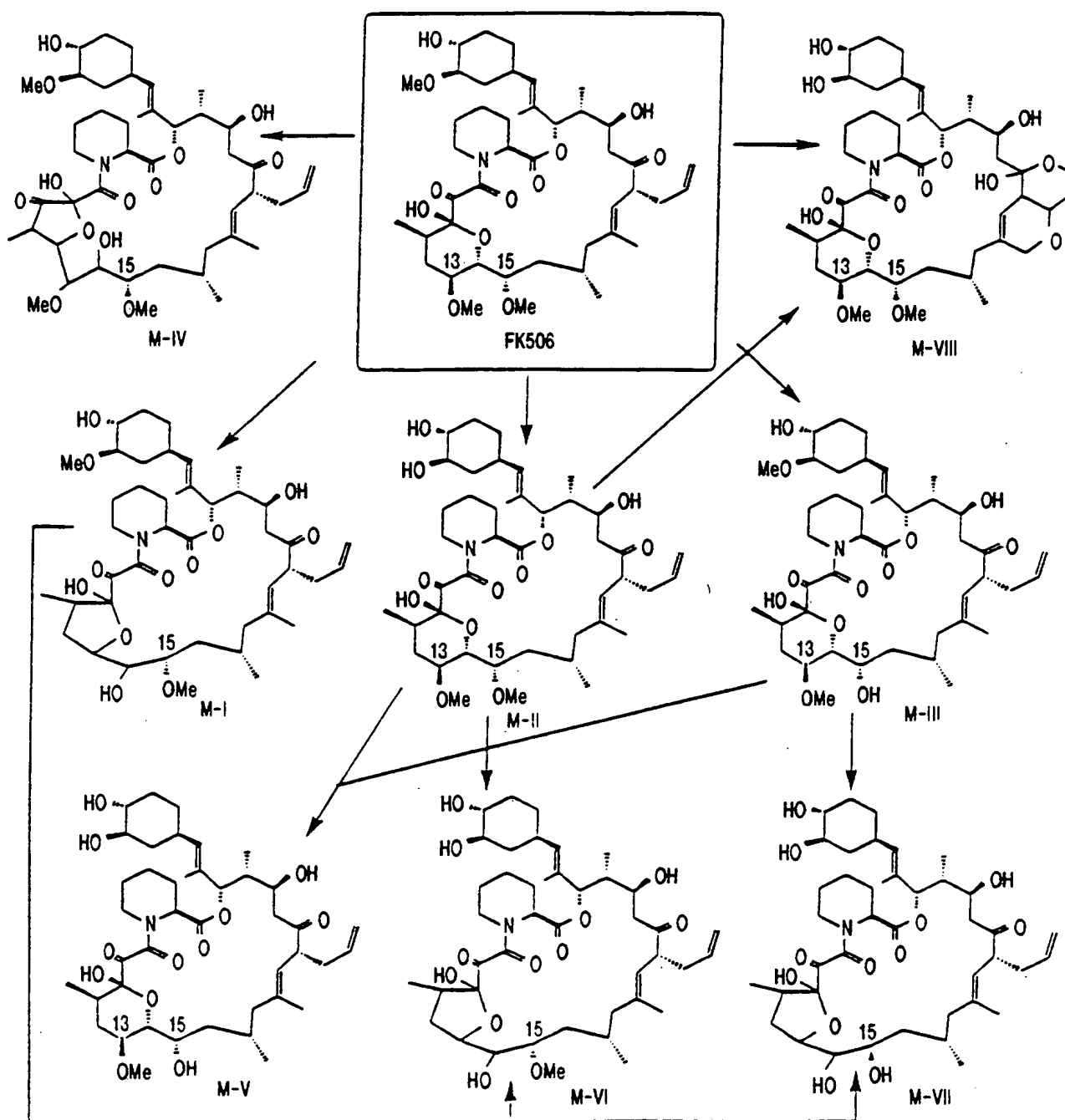


FIG. 6

SUBSTITUTE SHEET (RULE 26)

7 / 9

FIG. 7A



↓ linker insertion

FIG. 7B



↓ PCR amplification

FIG. 7C

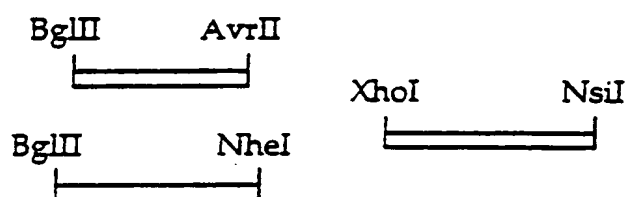
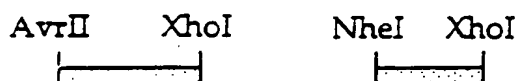
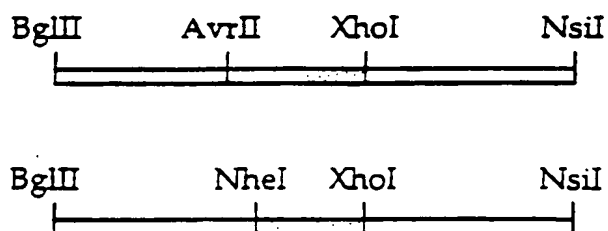


FIG. 7D



↓ ligation

FIG. 7E



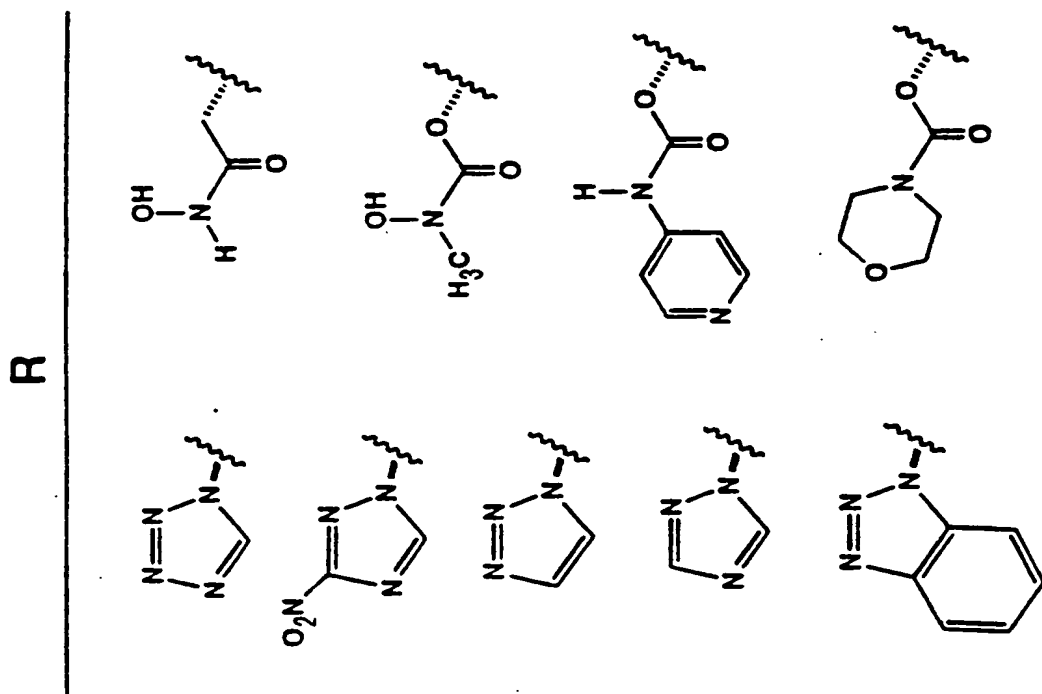
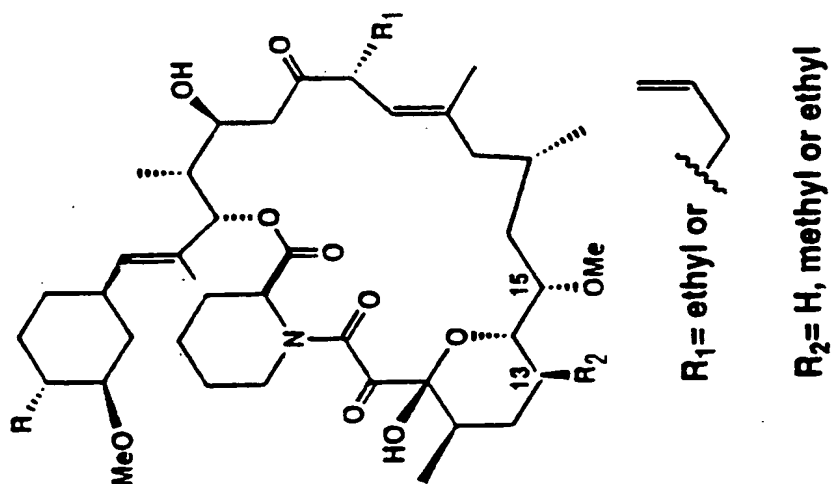


FIG. 8A



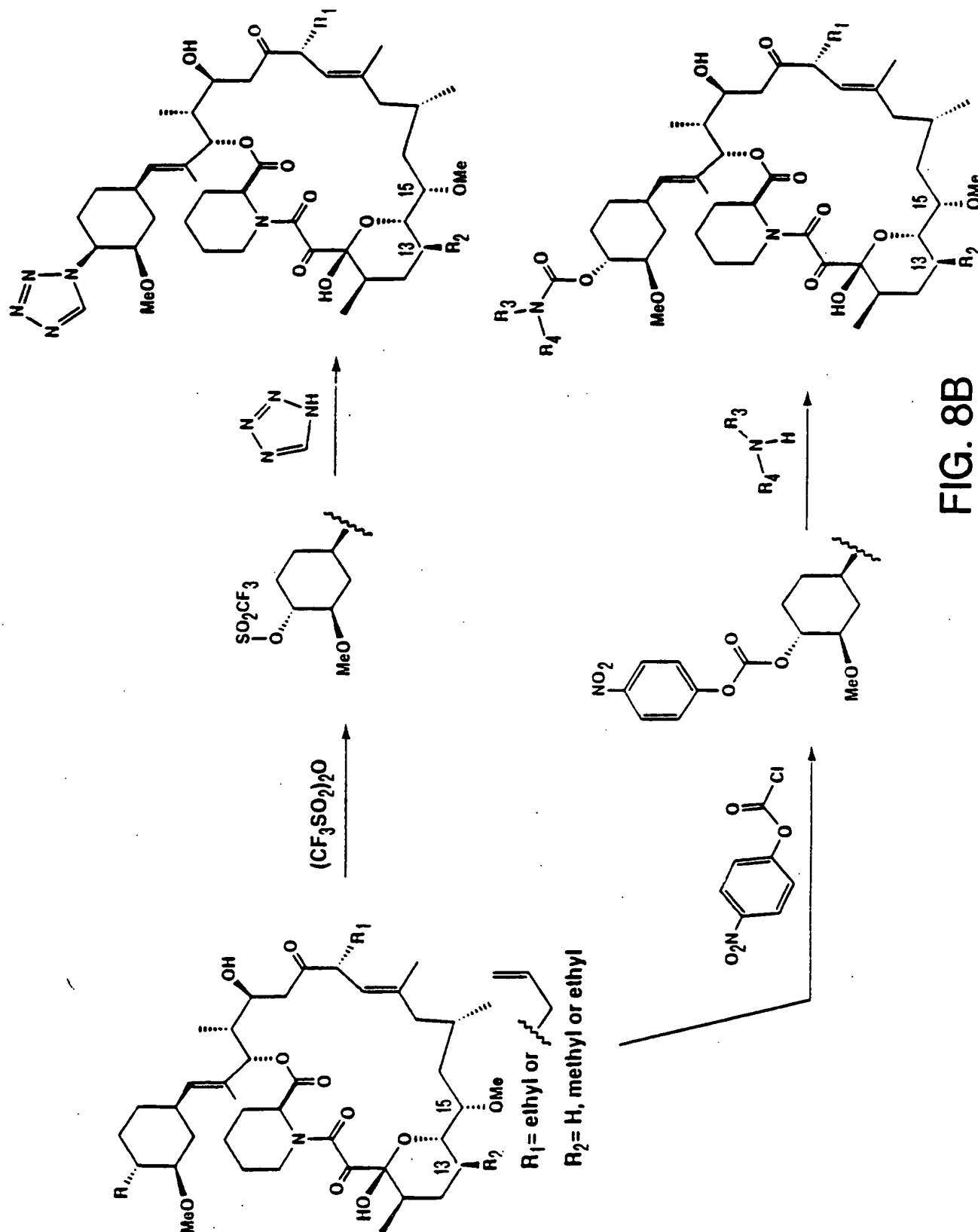


FIG. 8B

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

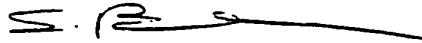
(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>22</u> , line <u>31-33</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) <div style="text-align: center;">10801 University Blvd Manassas, VA 22110-2209 USA</div>	
Date of deposit 20 September 1999	Accession Number PTA-727, PTA-728 and PTA-729
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
All designated States	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer</div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px;"><input checked="" type="checkbox"/> This sheet was received by the International Bureau on: 13 JUN 00</div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer S. B.</div>

Form PCT/RO/134 (July 1992)

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>22</u> , line <u>31-33</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) <div style="text-align: center;">10801 University Blvd Manassas, VA 22110-2209 USA</div>	
Date of deposit <div style="text-align: center;">20 September 1999</div>	Accession Number <div style="text-align: center;">PTA-726</div>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
<div style="text-align: center;">All designated States</div>	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer</div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px;"><input checked="" type="checkbox"/> This sheet was received by the International Bureau on: <div style="text-align: center; font-size: 1.2em;">13 JUN 00</div></div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer <div style="text-align: center;"></div></div>

Form PCT/RO/134 (July 1992)



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>7</sup> : C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)</p>	A3	<p>(11) International Publication Number: <b>WO 00/20601</b></p> <p>(43) International Publication Date: 13 April 2000 (13.04.00)</p>									
<p>(21) International Application Number: PCT/US99/22886</p> <p>(22) International Filing Date: 1 October 1999 (01.10.99)</p> <p>(30) Priority Data:</p> <table border="0"> <tr> <td>60/102,748</td> <td>2 October 1998 (02.10.98)</td> <td>US</td> </tr> <tr> <td>60/123,810</td> <td>11 March 1999 (11.03.99)</td> <td>US</td> </tr> <tr> <td>60/139,650</td> <td>17 June 1999 (17.06.99)</td> <td>US</td> </tr> </table> <p>(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altamir Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).</p>		60/102,748	2 October 1998 (02.10.98)	US	60/123,810	11 March 1999 (11.03.99)	US	60/139,650	17 June 1999 (17.06.99)	US	<p>(74) Agents: FAVORITO, Carolyn et al.; Morrison &amp; Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).</p> <p>(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report.</p> <p>(88) Date of publication of the international search report: 26 October 2000 (26.10.00)</p>
60/102,748	2 October 1998 (02.10.98)	US									
60/123,810	11 March 1999 (11.03.99)	US									
60/139,650	17 June 1999 (17.06.99)	US									
<p>(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR</p> <div data-bbox="415 1142 1252 1730"> </div> <p>(57) Abstract</p> <p>Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.</p>											

\*(Referred to in PCT Gazette No. 35/2000, Section II)

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/22886

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/52 C12N15/54 C12N15/62 C12N9/10 C12P17/18  
C12P19/32 C07D498/18 //(C07D498/18,311:00,273:00,211:00)

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12P C07D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, MEDLINE, STRAND, BIOSIS, EMBASE, CHEM ABS Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOTAMEDI H ET AL.: "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506" EUR. J. BIOCHEM., vol. 256, no. 3, 15 September 1998 (1998-09-15), pages 528-534, XP000906738 abstract figures 1,2 page 532, right-hand column, line 51 -page 533, left-hand column, line 18 --- -/--	12

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

27 July 2000

Date of mailing of the international search report

10/08/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22886

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	STASSI D L ET AL.: "Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering" PROC. NATL. ACAD. SCI. USA, vol. 95, no. 13, 23 June 1998 (1998-06-23), pages 7305-7309, XP002143632 abstract page 7308, left-hand column, line 4 -right-hand column, line 17 page 7309, left-hand column, line 24-40 ---	17
X	REYNOLDS K A ET AL.: "Rapamycin, FK506, and ascomycin -related compounds" DRUGS PHARM. SCI., vol. 82, 1997, pages 497-520, XP000906777 figure 3 page 502, line 7-25; figure 7 page 509-513, paragraph IV ---	18
X	DUMONT F J ET AL.: "The immunosuppressive and toxic effects of FK-506 are mechanistically related: pharmacology of a novel antagonist of FK-506 and rapamycin" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 176, no. 3, 1 September 1992 (1992-09-01), pages 751-760, XP000906781 cited in the application abstract; figure 1 ---	18
X	KAWAI M ET AL.: "Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues" FEBS LETTERS, vol. 316, no. 2, January 1993 (1993-01), pages 107-113, XP002143633 abstract scheme 1 table 1 ---	18
X	EP 0 323 042 A (FISONS PLC) 5 July 1989 (1989-07-05) example 13 ---	18
X	EP 0 356 399 A (SANDOZ AG ;SANDOZ AG (DE); SANDOZ AG (AT)) 28 February 1990 (1990-02-28) examples 2,3 ---	18
X	EP 0 463 690 A (MERCK & CO INC) 2 January 1992 (1992-01-02) example 3 ---	18
	-/--	

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22886

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
Y	MOTAMEDI H ET AL.: "Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520" J. BACTERIOLOGY, vol. 178, no. 17, July 1996 (1996-07), pages 5243-5248, XP002137077 abstract page 5245, left-hand column, line 1-3 figure 4	1-11
Y	--- MOTAMEDI H ET AL.: "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506" EUR. J. BIOCHEM., vol. 244, no. 1, 15 February 1997 (1997-02-15), pages 74-80, XP000906743 abstract page 79, left-hand column, line 26-35 page 75, left-hand column, line 31 -page 76, left-hand column, line 1	1-11
A	--- CHEN T S ET AL.: "Microbial transformation of immunosuppressive compounds. II. Specific desmethylation of 13-methoxy group of FK 506 and FR 9500520 by Actinomycete sp. ATCC 53828" J. ANTIBIOT., vol. 45, no. 4, April 1992 (1992-04), pages 577-580, XP002143634 figure 1	18-20
A	--- SHAFIEE A ET AL.: "Enzymatic synthesis and immunosuppressive activity of novel desmethylated immunomycins (ascomycins)" J. ANTIBIOT., vol. 46, no. 9, September 1993 (1993-09), pages 1397-1405, XP002143635 abstract	18,20
A	--- KHOSLA C: "Harnessing the biosynthetic potential of modular polyketide synthases" CHEMICAL REVIEWS, vol. 97, no. 7, 1997, pages 2577-2590, XP002130646 ISSN: 0009-2665 -----	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US-99/22886

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0323042 A	05-07-1989	AT 120466 T	15-04-1995
		AU 630866 B	12-11-1992
		AU 2822889 A	05-07-1989
		CA 1339128 A	29-07-1997
		CN 1033458 A	21-06-1989
		DE 3853477 D	04-05-1995
		DE 3853477 T	09-11-1995
		DK 387889 A	08-08-1989
		EP 0346427 A	20-12-1989
		ES 2071681 T	01-07-1995
		FI 90550 B	15-11-1993
		FI 930597 A	11-02-1993
		WO 8905304 A	15-06-1989
		IE 66163 B	13-12-1995
		IL 88629 A	12-04-1994
		JP 2502463 T	09-08-1990
		JP 2799208 B	17-09-1998
		NO 893166 A	04-08-1989
		NZ 227251 A	26-02-1990
		PT 89203 A, B	29-12-1989
		US 5376663 A	27-12-1994
		ZA 8809136 A	30-08-1989
EP 0356399 A	28-02-1990	AU 629563 B	08-10-1992
		AU 4024689 A	01-03-1990
		DK 418789 A	27-02-1990
		JP 2167287 A	27-06-1990
		NZ 230418 A	25-10-1991
		US 5011844 A	30-04-1991
		ZA 8906524 A	24-04-1991
EP 0463690 A	02-01-1992	CA 2044846 A	26-12-1991
		DE 69117221 D	28-03-1996
		DE 69117221 T	12-09-1996
		JP 4230288 A	19-08-1992
		US 5190950 A	02-03-1993
		US 5342935 A	30-08-1994

RECEIVED  
MAY 13 2002  
OPE/JCWS